DNA methylation, transcriptome and genetic copy number signatures of diffuse cerebral WHO grade II/III gliomas resolve cancer heterogeneity and development

Binder, H ; Willscher, E ; Loeffler-Wirth, H ; Hopp, L ; Jones, D T W ; Pfister, S M ; Kreuz, M ; et al ; Weller, Michael

Abstract: BACKGROUND Diffuse lower WHO grade II and III gliomas (LGG) are slowly progressing brain tumors, many of which eventually transform into a more aggressive type. LGG is characterized by widespread genetic and transcriptional heterogeneity, yet little is known about the heterogeneity of the DNA methylome, its function in tumor biology, coupling with the transcriptome and tumor microenvironment and its possible impact for tumor development. METHODS We here present novel DNA methylation data of an LGG-cohort collected in the German Glioma Network containing about 85% isocitrate dehydrogenase (IDH) mutated tumors and performed a combined bioinformatics analysis using patient-matched genome and transcriptome data. RESULTS Stratification of LGG based on gene expression and DNA-methylation provided four consensus subtypes. We characterized them in terms of genetic alterations, functional context, cellular composition, tumor microenvironment and their possible impact for treatment resistance and prognosis. Glioma with astrocytoma-resembling phenotypes constitute the largest fraction of nearly 60%. They revealed largest diversity and were divided into four expression and three methylation groups which only partly match each other thus reflecting largely decoupled expression and methylation patterns. We identified a novel G-protein coupled receptor and a cancer-related 'keratinization' methylation signature in in addition to the glioma-CpG island methylator phenotype (G-CIMP) signature. These different signatures overlap and combine in various ways giving rise to diverse methylation and expression patterns that shape the glioma phenotypes. The decrease of global methylation in astrocytoma-like LGG associates with higher WHO grade, age at diagnosis and inferior prognosis. We found analogies between astrocytoma-like LGG with grade IV IDH-wild type tumors regarding possible worsening of treatment resistance along a proneural-to-mesenchymal axis. Using gene signature-based inference we elucidated the impact of cellular composition of the tumors including immune cell bystanders such as macrophages. CONCLUSIONS Genomic, epigenomic and transcriptomic factors act in concert but partly also in a decoupled fashion what underpins the need for integrative, multidimensional stratification of LGG by combining these data on gene and cellular levels to delineate mechanisms of gene (de-)regulation and to enable better patient stratification and individualization of treatment.

DOI: https://doi.org/10.1186/s40478-019-0704-8

Originally published at:
Binder, H; Willscher, E; Loeffler-Wirth, H; Hopp, L; Jones, D T W; Pfister, S M; Kreuz, M; et al; Weller, Michael (2019). DNA methylation, transcriptome and genetic copy number signatures of diffuse cerebral WHO grade II/III gliomas resolve cancer heterogeneity and development. Acta Neuropathologica Communications, 7(1):59.
DOI: https://doi.org/10.1186/s40478-019-0704-8

DNA methylation, transcriptome and genetic copy number signatures of diffuse cerebral WHO grade II/III gliomas resolve cancer heterogeneity and development

H. Binder[#§1], E. Willscher[§1], H. Loeffler-Wirth[1], L. Hopp[1], D. T. W. Jones[2,3], S. M. Pfister[2,4,5], M. Kreuz[6], D. Gramatzki[7], E. Fortenbacher[1], B. Hentschel[6], M. Tatagiba[8], U. Herrlinger[9], H. Vatter[9], J. Matschke[10], M. Westphal[11], D. Krex[12], G. Schackert[12], J.C. Tonn[13], U. Schlegel[14], H.-J. Steiger[15], W. Wick[16,17], R. G. Weber[18], M. Weller[+7], M. Loeffler[+6]

1 Interdisciplinary Centre for Bioinformatics, Universität Leipzig, Härtelstr. 16–18, 04107 Leipzig, Germany

2 Hopp Children's Cancer Center Heidelberg (KiTZ), Im Neuenheimer Feld 430, 69120 Heidelberg, Germany

3 Pediatric Glioma Research Group, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

4 Division of Pediatric Neurooncology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

5 Department of Pediatric Oncology, Hematology and Immunology, Heidelberg University Hospital, Im Neuenheimer Feld 430, 69120 Heidelberg, Germany

6 Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Härtelstraße 16-18, 04107 Leipzig, Germany

7 Department of Neurology, University Hospital and University Zurich, Frauenklinikstrasse 26, 8091 Zurich, Switzerland

8 Clinic for Neurosurgery, Tübingen University Hospital, Hoppe-Seyler-Straße 3, 72076 Tübingen, Germany

9 Division of Clinical Neurooncology, Department of Neurology, University Hospital Bonn, Bonn, Germany

10 Institute of Neuropathologie, University Clinic Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany

11 Department of Neurosurgery, University Clinic Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany

12 Department of Neurosurgery, Technical University Dresden, Fetscherstraße 74, 01307 Dresden, Germany

13 Department of Neurosurgery, Ludwig Maximilians University Munich and German Cancer Consortium (DKTK), partner site Munich, Marchioninistraße 15, D-81377 Munich, Germany

14 Department of Neurology, University Hospital Knappschaftskrankenhaus Bochum-Langendreer, In der Schornau 23-25, 44892 Bochum, Germany

15 Clinic for Neurosurgery, University Düsseldorf, Moorenstr. 5, 40225 Düsseldorf, Germany

16 Clinical Cooperation Unit Neurooncology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

17 Neurology Clinic and National Center for Tumor Diseases, University Hospital Heidelberg, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany

18 Department of Human Genetics, Hannover Medical School, Carl-Neuberg-Str. 1, 30625 Hannover, Germany

To whom correspondence should be send
§ contributed equally
+ shared last authorship

List of e-mail addresses:

Hans Binder	binder@izbi.uni-leipzig.de
Edith Willscher	willscher@izbi.uni-leipzig.de
Henry Loeffler-Wirth	wirth@izbi.uni-leipzig.de
Lydia Hopp	hopp@izbi.uni-leipzig.de
David T. Jones	david.jones@kitz-heidelberg.de
Stefan Pfister	s.pfister@kitz-heidelberg.de
Markus Kreuz	mkreuz@imise.uni-leipzig.de
Dorothee Gramatzki	Dorothee.Gramatzki@usz.ch
Erik Fortenbacher	e.fortenbacher@gmx.de
Bettina Hentschel	bettina.hentschel@imise.uni-leipzig.de
Marcos Tatagiba	marcos.tatagiba@med.uni-tuebingen.de
Ullrich Herrlinger	Ulrich.Herrlinger@ukbonn.de
Hartmut Vatter	hartmut.vatter@ukb.uni-bonn.de
Jakob Matschke	matschke@uke.de
Manfred Westphal	westphal@uke.uni-hamburg.de
Dietmar Krex	Dietmar.Krex@uniklinikum-dresden.de
Gabriele Schackert	gabriele.schackert@uniklinikum-dresden.de
Joerg Christian Tonn	Joerg.Christian.Tonn@med.uni-muenchen.de
Uwe Schlegel	uwe.schlegel@kk-bochum.de
Hans-Jakob Steiger	Hans-Jakob.Steiger@med.uni-duesseldorf.de
Wolfgang Wick	wolfgang.wick@med.uni-heidelberg.de
Ruthild G. Weber	Weber.Ruthild@mh-hannover.de
Michael Weller	Michael.Weller@usz.ch
Markus Loeffler	markus.loeffler@imise.uni-leipzig.de

Abstract

Background: Diffuse lower WHO grade II and III gliomas (LGG) are slowly progressing brain tumors, many of which eventually transform into a more aggressive type. LGG is characterized by widespread genetic and transcriptional heterogeneity, yet little is known about the heterogeneity of the DNA methylome, its function in tumor biology, coupling with the transcriptome and tumor microenvironment and its possible impact for tumor development.

Methods: We here present novel DNA methylation data of an LGG-cohort collected in the German Glioma Network containing about 85% *isocitrate dehydrogenase (IDH)* mutated tumors and performed a combined bioinformatics analysis using patient-matched genome and transcriptome data.

Results: Stratification of LGG based on gene expression and DNA-methylation provided four consensus subtypes. We characterized them in terms of genetic alterations, functional context, cellular composition, tumor microenvironment and their possible impact for treatment resistance and prognosis. Glioma with astrocytoma-resembling phenotypes constitute the largest fraction of nearly 60%. They revealed largest diversity and were divided into four expression and three methylation groups which only partly match each other thus reflecting largely decoupled expression and methylation patterns. We identified a novel G-protein coupled receptor and a cancer-related 'keratinization' methylation signature in in addition to the glioma-CpG island methylator phenotype (G-CIMP) signature. These different signatures overlap and combine in various ways giving rise to diverse methylation and expression patterns that shape the glioma phenotypes. The decrease of global methylation in astrocytoma-like LGG associates with higher WHO grade, age at diagnosis and inferior prognosis. We found analogies between astrocytoma-like LGG with grade IV *IDH*-wild type tumors regarding possible worsening of treatment resistance along a proneural-to-mesenchymal axis. Using gene signature-based inference we elucidated the impact of cellular composition of the tumors including immune cell bystanders such as macrophages.

Conclusions: Genomic, epigenomic and transcriptomic factors act in concert but partly also in a decoupled fashion what underpins the need for integrative, multidimensional stratification of LGG by combining these data on gene and cellular levels to delineate mechanisms of gene (de-)regulation and to enable better patient stratification and individualization of treatment.

Key words: glioma, molecular subtypes, DNA methylation, epigenetics, astrocytoma, tumor microenvironment, cellular composition, prognosis

Background

Diffuse WHO grade II and III glioma (in short lower grade glioma, LGG) describe an almost fatal disease of young adults. These tumors share a more indolent course compared with high-grade IV gliomas (glioblastoma, GBM). Recent work has proposed a classification of glioma based mainly on two genetic markers, namely absence or presence of isocitrate dehydrogenase 1 and 2 (*IDH*) mutation and of codeletion of chromosome arms 1p and 19q (codel), overriding histology [1-5]. *IDH*-mut codel tumors with mostly oligodendroglial histology are associated with the best prognosis; *IDH*-mut non-codel tumors with mostly astrocytic histology are associated with intermediate outcome; and *IDH*-wt, with mostly higher WHO grade (III or IV) tumors are associated with poor prognosis [6, 7]. Besides genetic factors (DNA-)methylation has emerged an important regulator of gene transcription, and its role in tumorigenesis has become a topic of considerable interest [8, 9]. A number of studies have reported alterations of DNA methylation in gliomas [10-18].

IDH mutations occur early in gliomagenesis in the vast majority of WHO grade II and III gliomas. They change the function of the *IDH* enzymes, causing them to produce 2-hydroxyglutarate (2HG), an oncometabolite that represses DNA demethylation [19] and, in consequence, leads to genome wide DNA-hypermethylation subsumed as glioma-CpG island methylator phenotype (G-CIMP) [14]. Whole genome methylation studies have revealed that G-CIMP gliomas split into subgroups differing in the Chr. 1p/19q codeletion status and the total level of methylation [20] where decreased methlylation associates with worse survival and increased risk for recurrence [17] and possibly reflects a global DNA demethylation shift of progressing G-CIMP-tumors. Both, genetic and epigenetic events can drive progression of gliomas leading to nearly identical phylo(epi-)genetic relations [21]. Moreover, recent studies reported continuous phenotypic drifts along a proneuronal-to-mesenchymal axis in *IDH*-wild type GBM associated with increasing therapy resistance that contradict a major role of genetic aberrations as drivers of essential tumor characteristics such as resistance [22] and that are linked to drifts in DNA methylation [23] and the cellular composition of the tumor microenvironment [24]. We ask if similar mechanisms can be identified also in *IDH* mutated LGG.

In general, deregulation of cell functions in cancer is encoded in both the genome and epigenome which underscores the importance of epigenetic analyses in parallel to the discovery of transcriptomics and genetics. Current analyses have not yet clarified the relationships between the methylome and transcriptome. In LGGs about 84% of all cases carry *IDH*-mutation with about 35% of them carrying an additional Chr. 1p/19q-codeletion, which enables studying phenotypic variability of the transcriptomes and methylomes especially of these genomic strata.

Our previous expression profiling of grade II and III primary glioma from a prospective German Glioma Network (GGN) cohort revealed rich heterogeneity of their transcriptomes which were only partially linked to the genomic features [6]. For this study transcriptomic and genetic data of the 137 lower grade glioma samples from the GGN cohort were complemented by new (DNA-)methylation data of 122 matched tumors of the same cohort which enables a combined analysis aimed to study DNA methylation as a shaping factor of glioma heterogeneity. Here we perform molecular subtyping which has emerged as an important concept to describe glioma heterogeneity and to better understand the biology of this devastating disease. We show that genomic, transcriptomic and methylation data provide partly overlapping but also distinct molecular subgroups, suggesting that different omics-views provide complementary and partly independent information about modes of gene-regulation [25, 26] with potentially different prognostic and therapeutic relevance. We aimed at characterizing the functional context of these different modes with special emphasis on the cellular composition of the tumors and their microenvironment and on possible impact for tumor development from lower grade to higher grade gliomas.

Materials and Methods

Patients, tumors and data

The GGN study of WHO grade II or III gliomas (LGG) was described previously [6]. For this previous study, we had analyzed tumors of 137 patients by array-CGH, Affymetrix chip-based gene expression and candidate gene analyses (see [6] and Figure S 1). All tumors were subjected to central pathology review and classified according to the 2016 WHO classification of tumors of the central nervous system [3]. For the present study, molecular characterization was supplemented by array-based DNA methylation data (Illumina 450K arrays) of 122 patient-matched tumors of the GGN cohort (Table S 1).

Expression, CGH data and DNA methylation analyses

Expression and array CGH data were processed as described in [6]. For genome-wide assessment of DNA methylation glioma samples were arrayed using the Illumina HumanMethylation450 BeadChip according to the manufacturer's instructions at the DKFZ as described previously [16]. A verification set of WHO grade II and III gliomas was taken from the TCGA repository including gene expression and DNA methylation data (Table S 2). Gene expression data were corrected for background noise, calibrated, quantile-normalized and transformed into log10-scale, as described in [6]. CpG IDs were mapped to the promoter region of each gene ranging from 2kb upstream to 200bp downstream of the transcription start site using RefSeq mRNA annotation. DNA methylation beta-values of the respective CpGs were averaged to get one mean methylation beta-value for each gene promoter available. Genes located on Chr. X and Y were excluded from analyses. For an alternative analysis we also integrated CpG methylation over enhancer and gene body regions (see below).

Bioinformatics analysis

Gene expression and DNA methylation data were centralized and then analyzed after dimension reduction to metagenes using self-organizing map (SOM) machine learning [27]. As a result, each tumor tissue is characterized by the expression/methylation values of 2,500/900 metagenes. Downstream analysis tasks including class discovery, visualization and knowledge mining using gene set analysis were performed with the R-package 'oposSOM' [28]. Unsupervised class discovery of expression and methylation subgroups was performed independently in metagene space by using maximum spanning graph-partitioning [6] followed by iterative maximization of the sample similarity score until convergence as described before [29]. For gene set profiling we applied the gene set Z-score (GSZ) metrics to estimate the mean differential expression of the set genes in each sample compared with their mean expression levels averaged over all samples in units of the respective standard deviation [30]. We considered gene sets related to biological processes (BP), of the gene ontology (GO) classification, and standard literature sets and literature sets curated by our group [31]. Immune cell composition of the tumor biopsies were estimated from the expression data using the program CIBERSORT [32].

Results

WHO grade II/III gliomas split into eight expression and six methylation subtypes

Single data type class discovery of gene expression data of 137 WHO grade II/III gliomas and DNA methylation data of patient-matched samples all collected in the German Glioma Network (GGN) provided eight expression subtypes designated as E1 – E8 (E-groups) and six methylation subtypes (M1 – M6, M-groups, Figure 1A). The subtypes E1 and M1 nearly completely collect *IDH*-wild type tumors mostly with gains on Chr7 and losses of Chr10 representing genetic hallmarks of glioblastomas [6, 7, 33] (Figure 1B, Figure S 2 and Table S 4 and Table S 5 for sample counts and enrichment analysis). The subtypes E2 – E6 and M2 – M5 nearly exclusively contain *IDH* mutated tumors predominantly without codeletions on Chr1p and Chr19q as genetic hallmarks of astrocytomas while the subtypes E6 and M5 strongly enrich samples with a codeletion on Chr1p and Chr19q as a genetic hallmark of oligodendrogliomas [3]. Gains on Chr. 7 that are not paralleled by losses on Chr10 are frequently found in E4 and M3 (Figure S 2). A chromosome map of gene expression reveals dose-response effects of all these chromosomal defects (Figure S 3A). We find a relatively high number of aberrations in E2 and a relatively small one in E7/M6 and E5/M4 (Figure 1B and Figure S 2). Interestingly, a bimodal differential methylation pattern between M1 – M3 (reduced methylation) and M4 – M6 (increased methylation) is detected for the olfactory subgenome [34] which collects genes encoding G-protein coupled receptors (*GPCR*'s) especially on Chr11 (Figure S 3B). The E-groups do not show this clear separation into two entities although the amount of hypomethylated *GPCR*'s increases progressively from E7 to E1 (Figure S 3B). A similar bimodal methylation patterns is found for gene clusters encoding keratin intermediate filament proteins on Chr12 and Chr17 in the M groups (Figure S 3B). We compared the mean methylation levels of the promoter regions as used throughout this work with those of enhancer and gene body regions and found similar methylation patterns on average (Figure S 4), which suggests that aberrant methylation affects widespread genomic regions.

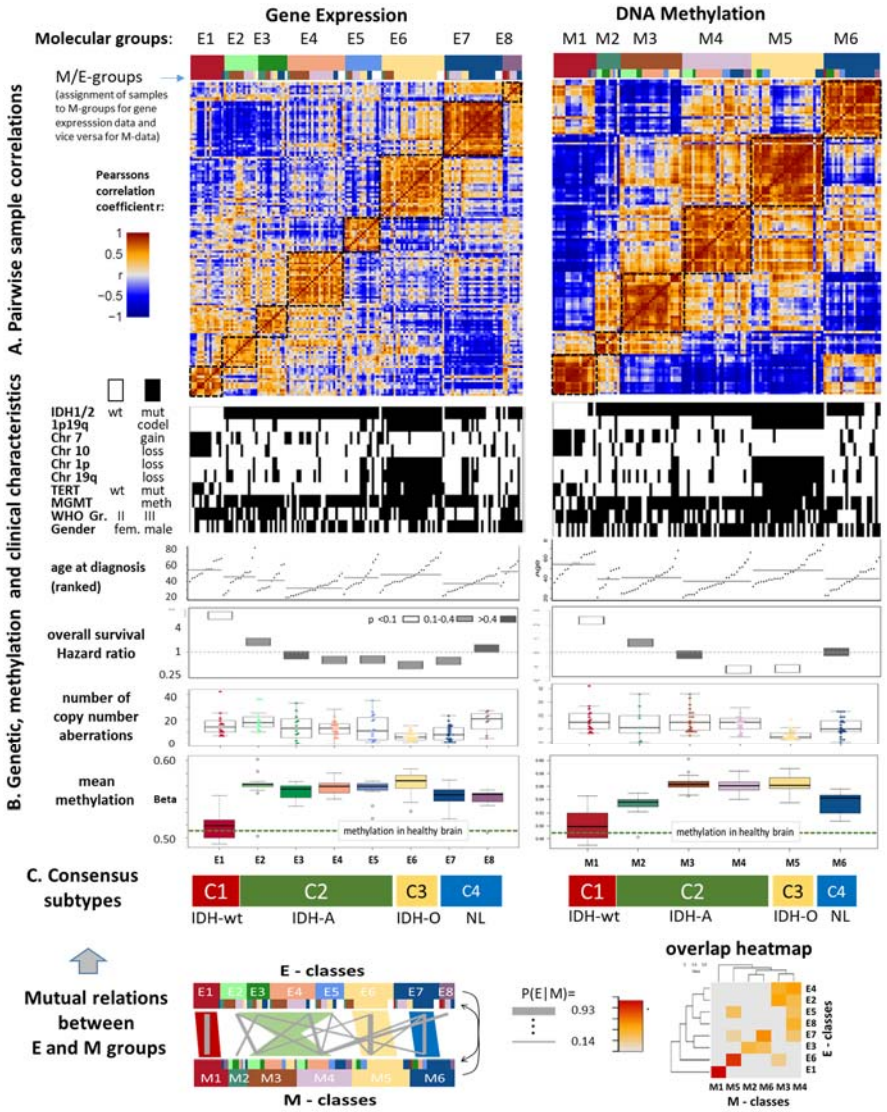


Figure 1: Characteristics of molecular subtypes of glioma. Samples were grouped into gene expression groups E1 – E8 (E-classes) or DNA methylation groups M1 – M6 (M-classes) using the sample expression and methylation data, respectively. A) The pairwise sample correlation heatmaps visualize the correlation coefficient between all pairwise combinations of sample-portraits. Intra-class similarities between samples are evident as brown quadratic areas along the diagonal while inter-class relations are seen either as brown or blue off-diagonal regions for positively and negatively correlated data landscapes, respectively. B) Genetic, methylation and clinical characteristics (see text). C) We sorted samples in each E-group according to their M-group membership and in each M-group according to their E-group membership to better recognize pattern due to methylation and expression effects, respectively (see the two color bars above the heatmap). The color code for molecular groups are used throughout the paper. Mutual relations between the E- and M-groups were estimated based on mutual memberships of the samples giving rise to four consensus subtypes C1- C4 which are characterized by IDH-wild type astrocytoma-like (IDH-wt), IDH-mutated astrocytoma-like (IDH-A) and oligodendroglioma-like (IDH-O) and a neuronal-like (NL) phenotypes, respectively.

Consensus subtypes assign to astrocytoma-like, oligodendroglioma-like and neural phenotypes

Detailed analysis of the distribution of samples among the E- and M-subtypes reveals large overlap of tumors and thus correspondence between E1 and M1, E6 and M5 and also between E7 and M6 (Figure 1C) while E2 – E5 intermix with M2 – M4 with partial correspondence between E3 and M2. Based on these results we define the consensus classes C1 – C4 where C1, C3 and C4 represent classes with almost one-to-one mutual correspondence between the expression and methylation subtypes. With a nearly exclusive content of *IDH*-wild type tumors in C1 (100% in E1 and 87.5% in M1) and of *IDH*-mutated *and* Chr1p/19q codeleted tumors in C3 (92% in E6 and 100% in M5) these subtypes show clear genetic characteristics that assign them to expression and methylation phenotypes of *IDH*-wild type astrocytoma-like (IDH-wt) and to oligodendrogioma-like (IDH-O) resemblance, respectively [3]. In contrast, C2 is a more heterogeneous group regarding the correspondence between the E- and M-classes. It collects predominantly *IDH* mutated tumors (more than 97 % in C2) almost always without Chr1p/19q codeletions (85% for E-groups and 90% for M-groups) and without alterations on Chr7 and Chr10 (95%) (see also Table S 4) which assigns C2 to gliomas of *IDH*-mutant astrocytoma-like resemblance (IDH-A) [3]. Nevertheless, a minority of about 15% of all *IDH*-mutant and Chr.1p/19q codeleted oligodendrogliomas are in C2 (12.5% in the E-groups and 17% in the M-groups) mostly because of a decreased methylation level of the *GPCR* subsumed in the olfactory subgenome that contrasts them compared with the majority of 67% oligodendrogliomas in C3/IDH-O (60% in E6 and 75% in M5) and also in C4 (15% / 8%, Figure S 3B). The consensus subtype C4 collecting E7, E8 and M6 constitutes mixtures of tumors with genetic characteristics present in all remaining subtypes. We assign specimens with reduced tumor cell content to C4 based on the observations that the mean number of copy number aberrations on Chr7 and Chr10 in E1 and on Chr1p and Chr19q in E6 is reduced for samples in E7, respectively (Figure S 5). Additionally, C4 shows a healthy brain functional context, e.g. related to synaptic transmission (see below). Hence, the expression and methylation landscapes of the glioma subtypes are shaped in first instance by the underlying key genetic defects in agreement with a recent classification of LGGs [4]. However, we also found a large degree of inter-tumoral heterogeneity of expression and methylation phenotypes that considerably modulates this genetic picture as illustrated by means of sample-similarity nets based either on the gene expression or on the methylation data (Figure S 7). This uncertainty obviously results, among other factors (such as tumor purity and composition), from the multidimensional nature of the transcriptomes and methylomes, e.g. from the combination of different G-CIMP- and *GPCR*-methylation patterns, from the lack of a clear-cut one-to-one relation between many of these phenotype-dimensions and the underlying genotypes.

The subtypes differ in overall promoter methylation, WHO grade and prognosis

Next, we compared the mean absolute promoter methylation level averaged over all genes and samples of each subtype (Figure 1). It is low in C1 (*IDH*-wt) and high in C3 (IDH-O) and also C2 (IDH-A), as expected, because these predominantly *IDH*-mutated tumors in C2 and C3 form the CpG hypermethylation phenotype (G-CIMP) [14]. The degree of hypermethylation in M2 is reduced compared with the other *IDH* mutated tumors in C2 and C3 while promoter methylation is on intermediate level in C4 collecting a mixture of *IDH*-mutated (64%) and *IDH*-wt (36%) tumors. Interestingly, the mean methylation level of the subtypes inversely relates to their overall survival hazard ratio (Figure S 6). Worst prognosis of *IDH* wild type LGG compared with *IDH* mutated *and* Chr1p/19q codeleted (best prognosis) and non-codeleted ones (intermediate prognosis) was reported previously [6]. Interestingly, we find a similar, however more subtle inverse trend between

methylation and HR in the E- and M-groups collected in C2 (IDH-A) that associates with the accumulation of WHO grade II astrocytic tumors in E4 (58%) and E5 (60%, decreased HR and increased methylation, Table S 4) while grade III tumors accumulate in E3 (100%) and E2 (71%, increased HR and decreased methylation). Enrichment of higher tumor grade III is also found in M2 (78%) and M3 (70%). It is associated with worse prognosis and decreased methylation (Figure S 6). Taken together, our data suggest associations between decaying methylation, increasing WHO grade and HR in *IDH* mutated astrocytoma-like tumors (IDH-A).

Verification using TCGA data and comparison with previous signatures of gliomas

The E- and M-subtypes found here were confirmed (except E5) in more than 270 LGGs taken from The Cancer Genome Atlas (TCGA) using a guided SOM-extension method that combines the GGN- and TCGA-data and enables their joint analysis [29] (Figure S 8, Figure S 9). Moreover, we selected a series of GBM and lower-grade glioma (LGG) signature gene sets of previous classification schemes and compared them with the subtypes identified herein (Figure 2 and Table S 6). We found correspondence between our subtypes E1 and partly E3 and signatures of the classical (CL) and mesenchymal (ME) expression subtypes of grade IV gliomas [35], of the pre-glioblastoma (PG) subtype of LGG [36] and of hypermethylated genes of the G-CIMP-phenotype [14]. Signature genes of proneural (PN) GBM [35] and of early-progenitor-like (EPL) LGG [36] show similarities with C2 (IDH-A), with subtle differences between E2, E3 and E4, while neuronal GBM (NL) and healthy brain signatures match to C4 and partly C3 (IDH-O). Interestingly, E3 collects *IDH*-mutated glioma with an inflammatory, mesenchymal-like expression signature. The expression level of most of these signatures sharply change between the E-groups which indicates correspondence between our current classification and those previously described. The analysis of gene sets derived from methylation studies provides analogous results where, e.g., hypermethylation signatures in LGG [13] and the *IDH* subtype of GBM [16] largely agree with the G-CIMP-profile [14] that shows hypermethylation in M2 - M5 (Figure 2). Oligodendroglial glioma reveal a modified G-CIMP-profile (G-CIMP-O) with enhanced methylation in M5 that closely resembles the RTKII profile [16]. Recent CpG-level marker sets confirm the G-CIMP and GCIMP-O profiles [37] (Figure S 11). Interestingly, methylation signatures of fetal and adult healthy brain [38] indicated strong similarity with the *GPCR*-signature meaning that the respective genes markedly lose methylation in gliomas, especially in M1 – M3. Overall, methylation signatures from previous studies including those of grade IV GBM [16] indicate similar underlying expression and methylation patterns. Accordingly, the consensus subtype C1 (IDH-wt) possesses pre-GBM (PG) characteristics, C2 (IDH-A) and C3 (IDH-O) are proneural-like (PN) tumors (with E3 showing more mesenchymal-like characteristics) and C4 represents a neural-like (NL) subtype with mixed genetic characteristics of gliomas and expression properties partly resembling those of healthy brain in agreement with [24]. We also compared our subtypes with the epigenetic classes of Ceccarelli et al. [20] making use of an extended GGN data set (Figure S 12). Accordingly, M2 tumors reveal resemblance with the GCIMP-low and M3 – M4 tumors with the GCIMP-high classes of Ceccarelli et al. while C1 tumors can be assigned to CL-like (8 cases) and ME-like (6 cases) glioma. Most interestingly, the E3 tumors reveal characteristics of pilocytic astrocytomas (PA) which was detected by comparison with the expression patterns of 16 PA samples collected in the GGN. PA-resemblance was established for IDH-wt gliomas by Ceccarelli and colleagues but not for IDH-mut LGG. In summary, almost all E- and M- subtypes could be verified in an independent dataset and by previous glioma signatures where our approach stratifies IDH-mut astrocytomas (C2) in a novel way into three methylation and four expression subtypes which only partly match each to another.

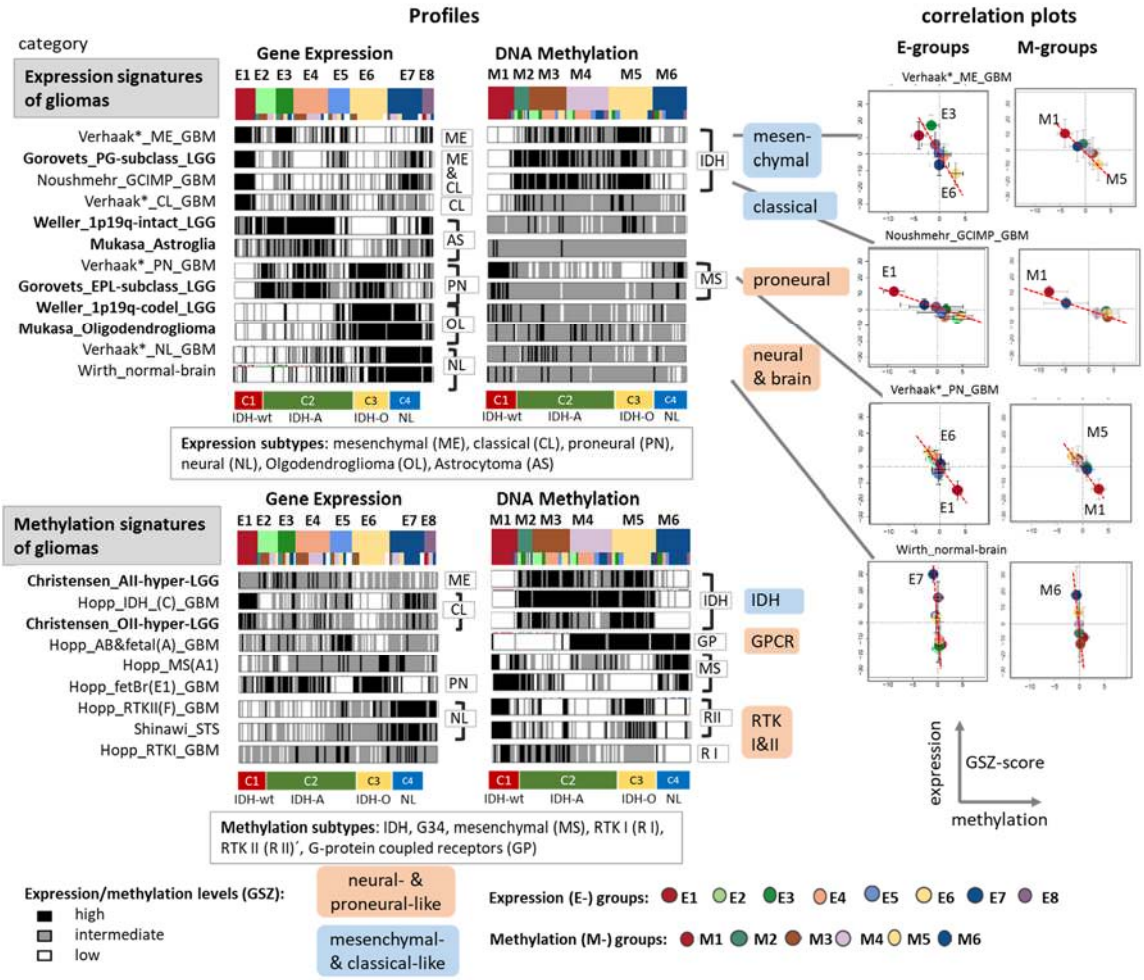


Figure 2: Gene set analysis associates the E- and M-subtypes with previous glioma expression and methylation signatures (see Table S 6 for details). The expression and methylation levels of the signature sets are shown as bar-code profiles where each bar refers to one sample. Correlation plots between expression and methylation levels in GSZ-scale reflect predominantly repressive effects of promoter methylation on the expression of the downstream genes (right part).

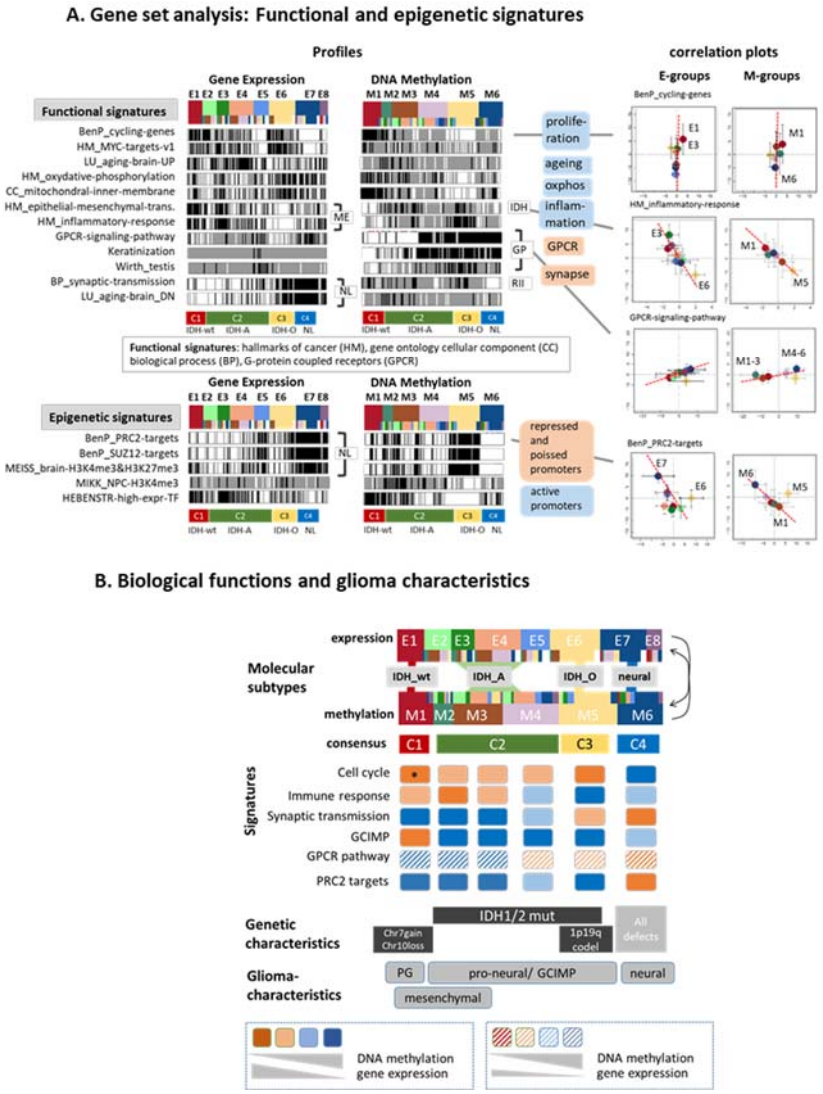


Figure 3: Gene set analysis of functional and epigenetic signatures: A) Bar-code profiles of expression and methylation levels of functional and epigenetic signatures and the correlation plots of subtype averaged values (see legend of Figure 2). B) Schematic overview about the basic functional, genetic and glioma characteristics extracted from the gene-signature analysis.

Functional context and epigenetic signatures

Next we analyzed the functional context of the E- and M-subtypes of our data set (Figure 3 and Figure S 13- Figure S 15). Gene signatures reflecting highly proliferating cells and high levels of oxphos metabolism are strongly expressed in E1 and E6 but weakly expressed in E7 which instead shows activated cell functions of healthy brain such as synaptic transmission. Inflammatory responses and a signature of epithelial-mesenchymal transition (EMT) were high in E3 and to a lesser degree observed in E1 but almost deactivated in C3 (IDH-O). Profiling of a series of metabolic gene sets confirms high oxphos and mitochondrial transcriptional activity in C3 paralleled by decreased glycolysis (Figure S 14) while C2 (IDH-A) is characterized by gained methylation and decreased expression of genes related to fatty acid metabolism, oxphos and mitochondrial functions. Interestingly, E2 seems metabolically deactivated throughout all processes considered while C1 (IDH-wt) shows the opposite trend.

Gene signatures of the ageing brain suggest parallels with inflammatory signatures upregulated in E3/E4. The methylation profiles of all these signatures show mostly anti-correlated patterns compared with the expression profiles (see the right 'methylation' part of Figure 3A and Figure S 13). It indicates that promoter methylation predominantly represses transcription of the respective downstream gene. Gene sets estimating the activity of G-protein coupled receptors (*GPCR*), and of keratinization both show binary methylation patterns with low levels in M1 – M3 and higher levels in M4 – M6 and correlated expression with the inflammatory signature and anti-correlated expression with the signature of synaptic transmission. Signatures related to epidermal cell differentiation and keratinization are prone to hypo-methylation also in other cancers [39]. They are found to tune the balance between stemness and somatic functions [40], to promote EMT-like processes [41] and also can serve as prognostic markers in epithelial cancers [42]. Testis-specific genes are overexpressed in E5. This phenomenon is observed also in other cancers [43] where so-called cancer testis (CT) genes often encode antigens that are thought to be immunogenic in gliomas and particularly in cancer stem cells [44-46].

Interestingly, also signatures with impact for epigenetic mechanisms of gene regulation reflect pronounced subtype-specific differences (Figure 3A, part below). Particularly, H3K4me3 marked genes in active promoters of neural progenitor cells (NPC) [47] and transcription factors (TF) associated with high expression levels in a wide collection of cells [48] show low methylation in M5 contrasted by high methylation in M1 and partly in M2 and M3 and thus similar trends as observed for the signatures related to highly proliferating cells and *MYC* targets discussed above. This seems plausible for M5 because highly proliferative cells require promoters activated by demethylation while activation of proliferation genes in M1 requires another mechanism. In contrast, hypermethylation in C3 (IDH-O), and to a less degree in C1 (IDH-wt) and C2 (IDH-A), is observed for targets of the polycomb repressive complex 2 (*PRC2*) in de-differentiated tumor cells [49], for related compounds such as *SUZ12* and *EED* targets and for bivalently H3K4me3 and H3K27me3 marked genes in poised promoter states that are enriched in tumor suppressors [50]. Their suppression via hypermethylation promotes cancer development in gliomas and in other cancer entities [38]. The respective expression and methylation profiles closely resemble those of healthy brain and synaptic transmission thus suggesting their suppression by epigenetics in gliomas. It is known that *PRC2* is required for neuron specification during differentiation and for suppression of a transcriptional program that is detrimental to adult neuron function and which in case of *PRC2* deficiency leads to neurodegeneration via de-repression of bivalent *PRC2* target genes [51, 52]. An analysis of genes in a set of defined chromatin states [53] determined in healthy fetal and adult brain tissues representing different states of brain development [54] further supports the view that suppressor-mechanisms in cellular programs are related to brain development and that genes in repressed states with impact for brain differentiation become hypermethylated in G-CIMP-subtypes and especially in C3 (Figure S 16).

Detailed functional analysis reveals anti-concerted alterations of expression and methylation, which associate with transcriptional activation of cell cycle related biological processes and the decay of neuronal processes such as synaptic transmission especially in C1 and C3 and also with changes of inflammatory characteristics in C2 (Figure S 17). Overall we identified three combined expression-methylation patterns (Figure S 18), namely (i) activating modes were related to proliferation and show increased expression which however associates either with increased (C1/IDH-wt) or decreased (C3/IDH-A) methylation reflecting different driving mechanisms; (ii) deactivating modes which combine decreased expression and increased methylation in all subtypes associated with functions such as synaptic transmission; and (iii) functions related to immune response also showing anti-correlated changes between expression and methylation but an activating effect in C1 and especially E3 and deactivating effect in C3. Hence, degeneration of healthy brain functions in all subtypes,

activated proliferation in C3 (IDH-O) and partly inflammation in E3 seem to be affected by anti-correlated DNA-promoter-methylation changes. In summary, the subtypes were characterized by combined alterations of the methylation and expression levels of genes from cellular programs such as proliferation, energy metabolism, immune response and synaptic transmission which associate with repressed and poised chromatin states in healthy brain and their subtype-specific remodeling in gliomas (Figure 3B).

Reference to single cell signatures disentangles glioma cell types

Gliomas are composed of neoplastic and non-neoplastic cells, each of which individually potentially contribute to cancer formation, progression and response to treatment [24, 55]. Bulk expression and methylation profiles as analyzed in this work average the diverse cell signatures within each tumor, thereby potentially masking critical differences and providing limited insight into cancer cell programs and the effect of the tumor microenvironment (TME). To disentangle this heterogeneity on cellular and TME-levels, we evaluated the expression and methylation degree of a collection of gene signatures taken from recent single cell RNAseq experiments on gliomas [56] in our data (Figure 4A). We find that C2 gliomas were characterized by relatively high expression levels of benign astrocytes (astro-program, especially in E4), malignant astrocyte-like cells (*IDH*-A signature) and of microglia/macrophages (especially in E3) which all confirm the astrocyte-like phenotype of C2. On the other hand, these signatures are all low in C3 tumors which instead show activated expression of oligodendrocyte-like cells (oligo-program and *IDH*-O signatures), of stemness and of neuronal signature genes where the latter ones are also high in C4 (neural subtype). The expression characteristics associate with almost mirror symmetrical methylation profiles showing either G-CIMP- or anti-G-CIMP characteristics, thus again suggesting regulatory effects of gene promoter methylation on downstream gene expression. A more detailed analysis indicates anti-correlated expression and methylation patterns of the malignant *IDH*-A and *IDH*-O dimensions suggesting that neoplastic transformations in *IDH*-O and *IDH*-A cells are driven by de-methylation of the respective signature genes while cell cycle and microglia/macrophage signatures increase and neuronal, healthy astro- and oligo-program signatures decline with increasing grade (Figure S 19). Overall, C3 tumors share closer similarities with healthy brain functions than C2-gliomas. C2-tumors instead show enhanced expression of macrophages/microglia signatures where microglia are crucial immune cells of the central nervous system and serve as tissue-resident macrophages of the brain [57]. On the other hand, both, C2 (IDH-A) and C3 (IDH-O) tumors are more proliferative compared with neuronal ones (C4). A higher amount of microglia/macrophage cells in astrocytoma and an increasing amount of proliferating cells is known to be a hallmark of higher grade gliomas [55]. In summary, the single cell characteristics reflect the variability of the composition of the tumors regarding healthy and benign astrocyte- and oligodendrocyte-like cells, microglia/macrophage and proliferative stem cell-like constituents in the bulk specimens studied.

mRNA inferred immune cell components

To further characterize the TME we employed CIBERSORT [32], a computational cell-type deconvolution method that decomposes the glioma-associated stroma into its immunological cellular components using cell-type related expression signatures (Figure 4B and Figure S 20). We find that M2-macrophages are highly abundant in the astrocytic groups E1 – E4 with highest levels in E3 and E2 opposed by a reduction in monocytes. M2-macrophages play a pro-tumoral and anti-inflammatory role in brain cancer; they promote tissue remodeling and tumor growth [57, 58], particularly in glioma [59], and associate with resistance to radiotherapy in mesenchymal glioblastoma [24]. In contrast, anti-

tumoral and pro-inflammatory M1-macrophages are almost absent in all grade LGG subtypes except for the pre-GBM subtype E1. Beyond a dual M1/M2 polarization status, a continuum between M1 and M2 polarization seems to exists in glioma [60] including *IDH*-mut tumors [56] which provides a possible interpretation of the increasing M2 percentage in C2 from E5 to E3. It has been hypothesized that the most aggressive and invasive cells in GBM are neoplastic macrophages arising in fusion hybrids between neoplastic stem cells and macrophages/microglia [61]. The high M2-macrophage abundance in astrocytic gliomas is paralleled by relatively large percentages of neutrophils while increased abundance of m1-macrophages in E1 is accompanied by CD4-resting memory T-cells. The amount of tumor infiltrating CD4+ leukocytes in glioblastoma correlates with tumor progression and presumably relates to tumor angiogenesis [62, 63]. We also found that activated mast cells are relatively abundant in virtually all groups (especially in E4, E5 and E7) except E1. Mast cells were shown to become recruited and 'educated' by glioma cells in a glioma grade-dependent manner to reduce stemness, decrease proliferation and migration to induce differentiation of glioma cells [64]. This mechanism seems to apply to early tumor stages of *IDH*-mutated astrocytoma-like gliomas (C2). Interestingly, regulatory T-cells (Tregs) show increased percentage in the oligodendroglioma-like subtype C3, the subtype with lowest immune and inflammatory characteristics, which is in correspondence with the immunosuppressive role of Tregs in glioma [65]. We also make use of immune cell gene signatures taken from [66] to compare their expression and methylation levels (Figure S 20). We find that most of them show high expression especially in E3 and E1 reflecting their accumulation in higher grade astrocytoma. These expression profiles are mostly paralleled by G-CIMP and especially G-CIMP-O methylation profiles which suggest deactivation of immune cell activities in C3 by DNA methylation. Interestingly, the methylation profile of T-cells resembles that of *GPCR*, which suggests a cell specific relation between DNA methylation and gene expression. Hence, the changes of methylation observed originate from both glioma and immune cells, which suggests coupled epigenetic mechanisms during tumor development. Note that DNA methylation in glioma bulk samples was found to be predictive for immune cell infiltration [23]. In summary, digital immune cell deconvolution of the transcriptome reveals that m2-macrophages were enriched in higher grade astrocytomas (E1- E3) while activated mast cells are more abundant in the neuronal subtype (C4), in lower grade astrocytomas (E4- E5) and in oligodendrogliomas (E3) together with immunosuppressive Tregs. Hence, the TME is characterized by marked variations of the immune cell composition that overlays with methylation changes of their genomes which suggests an epigenetically-mediated interplay between development of tumor cells and immune cells in the TME.

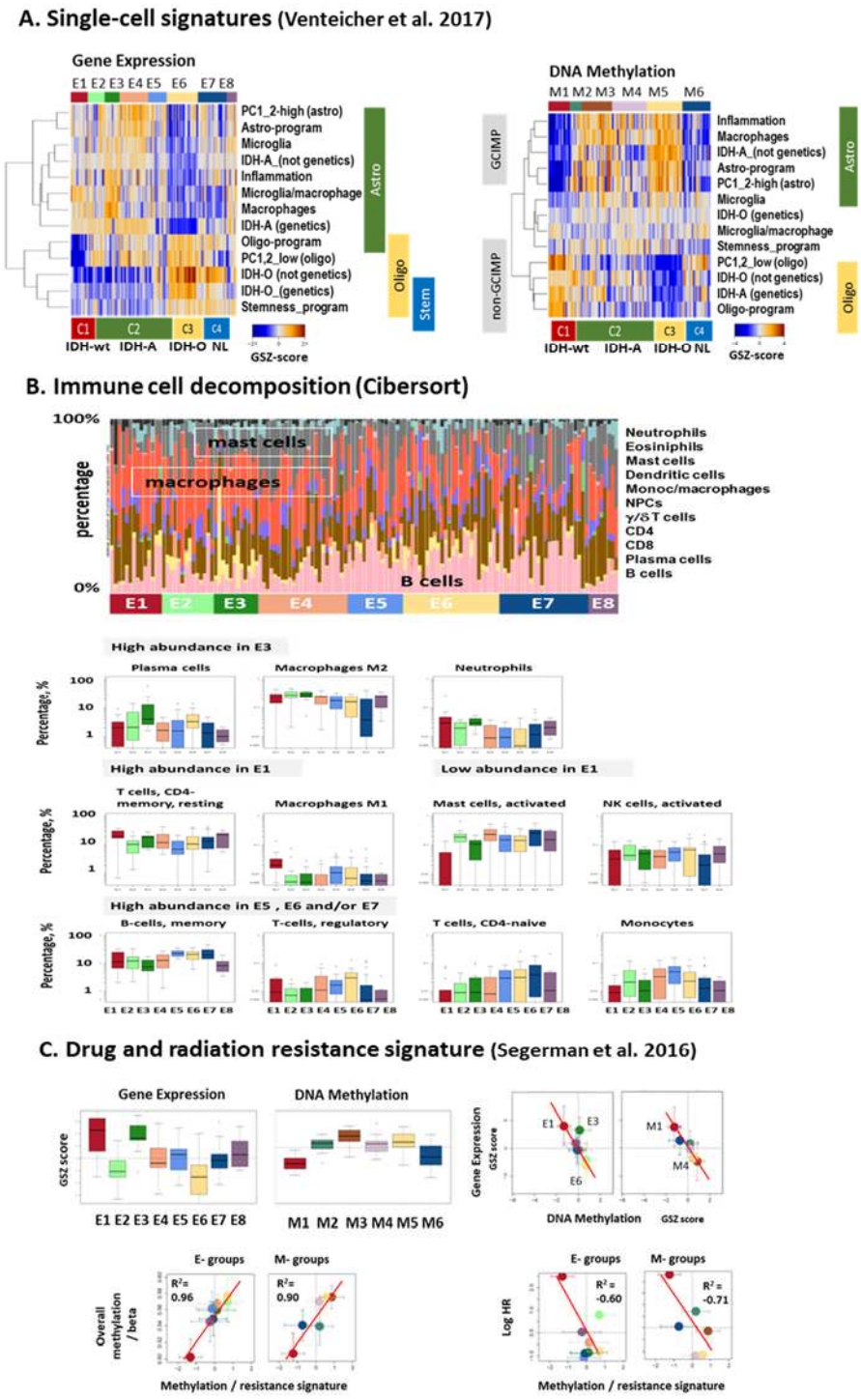


Figure 4: Cell type, micro-environmental immune cell and treatment-resistance characteristics. A) Heatmaps of expression and methylation levels of single-cell signatures taken from [56] reveal subtype-specific activation of astrocyte-, oligodendrocyte- and stem cell-like characteristics. B) Digital immune cell-type decomposition of glioma transcriptomes using CIBERSORT [32] (see Figure S 20 for the full set of cells considered) on sample (above) and mean subtype levels for selected leukocyte cells across the expression subtypes. C) The boxplots of expression and methylation levels of a transcriptomic drug and radiation resistance signature containing 50 genes [22] suggest largest resistance effects in E3 and E1. Expression and methylation levels of the subgroups anti-correlate (right part).

Treatment resistance and senescence signatures associate with methylation

Next, we studied a 50-gene multi-therapy resistance signature, which reflects a continuum of cell phenotypes with increasing resistance against chemo- and radiotherapy paralleled by a proneural-to-mesenchymal shift of their transcriptomes [22]. In our data, we find a profile of this signature showing highest expression in E3 (C2) and C1 and lowest in C3 (Figure 4C), thus suggesting a gradient of treatment resistance from oligodendroglioma-like to astrocytoma-like tumors with inflammatory characteristics of the TME. The resistance signature resembles the profiles of inflammatory and EMT functional signatures (Figure 3A) and that of the mesenchymal GBM-subtype (Figure 2A) in our data. Tumors of the latter type indeed showed enhanced treatment resistance [22-24]. The methylation profile of the resistance signature reflects G-CIMP characteristics and anti-correlates with the respective expression levels, which suggest a methylation-driven repression mechanism. Interestingly, the methylation profile of the resistance signature strongly correlates with the total methylation level of the tumors (R^2>0.9), which suggests that treatment resistance associates with overall methylation of the tumors. Moreover, the methylation profile of the signature anti-correlates with the HR-profile (R^2<-0.7, compare also with Figure 1) showing that worsening of prognosis associates with de-methylation, a trend that applies particularly also to the astrocytoma-like *IDH*-mutated tumors (C2).

A recent model of glioma progression suggests that increased senescence bypass mechanisms proceed in parallel with tumor development and the formation of a pro-inflammatory microenvironment at later phases [67]. We, therefore, studied a signature of genes that contribute to senescence bypass mechanisms by promoter-hypermethylation during aging and tumorigenesis and which associate with cancer risk [68]. These genes become increasingly deactivated in the tumors of the E-groups from E7 to E1, i.e. along the neuronal- pro-neural to mesenchymal axis (Figure S 22). Their senescence profile resembles those of the PRC2-targets, RTKII-characteristics, ageing and healthy brain signatures, while the methylation profiles of the two latter signatures differ from the former ones regarding methylation of Chr1p/19q-codeleted tumors in C3. Particularly, these tumors show increased methylation of senescence genes accompanied by demethylation and transcriptional upregulation of genes involved in oxphos-metabolism (Figure 3 and Figure S 14) and/or deactivated inflammatory response. It is assumed that Chr.1p/19q-codeleted gliomas (C3) bypass senescence by other mechanisms than Chr1p/19q-non-codeleted tumors [67]. Overall, the LGG-subtypes group along a therapy-resistance signature suggesting that resistance is mediated by epigenetics and an inflammatory TME along the proneural mesenchymal-like axis also in LGG. Moreover, astrocytoma-like tumors in C2 seem to develop along this axis as indicated by progressive activation of senescence bypass mechanisms.

Discussion

Heterogeneity of WHO grade II and III gliomas

Our multi-platform transcriptome-methylome-genome study revealed a large molecular heterogeneity of adult diffuse gliomas of WHO grades II and III: we identified eight expression and six methylation subtypes and characterized them in terms of genetic aberrations, functional context, cellular composition, tumor microenvironment and their possible impact on treatment resistance and prognosis as illustrated in the summary scheme in Figure 5. The expression and methylation patterns of the glioma subtypes are shaped by the underlying key genetic defects in agreement with recent classifications of LGG [4, 5, 20]. Overall, we identified three consensus subtypes C1-C3 that were assigned as *IDH*-wt and *IDH*-mut astrocytoma-like and oligodendroglioma-like phenotypes according to their dominating genetic status in terms of the *IDH* mutation and Chr. 1p/19q codeletion. These genetic aberrations are assumed to act as early events of tumorigenesis [69] (see left part of Figure

5A). A fourth, neuronal subtype (C4) collects specimen with reduced tumor cell content and served as reference partly resembling characteristics of healthy brain. However, our subtypes reflect also a large variability of expression and methylation phenotypes that do not match the genetic hallmarks in a one-to-one fashion. For example, 25-40% of all *IDH*-mut and 1p/19q-codel tumors were not assigned to the oligodendroglioma-like subtype (C3) but rather resemble the astrocytoma-like (C2) or neuronal (C4) types by a series of features. This heterogeneity results, among other possible factors, from the multidimensional nature of the transcriptomes and methylomes of the tumors. Each of their expression and methylation landscapes can be interpreted as a superposition of different expression and methylation patterns, which associate with specific cellular and micro-environmental states, and which obviously lack a clear-cut relation with respect to the underlying genotypes. The astrocytoma like gliomas constitute the largest fraction of nearly 60% of all LGG studied. They revealed the largest diversity and were divided into four expression (E2-E5) and three methylation (M2-M4) subtypes, which only partly match each other, thus reflecting partly decoupled expression and methylation patterns due to different possible interaction mechanisms [25, 26]. Particularly, decoupling between transcription and methylation can be rationalized in terms of independent regulation mechanisms of transcription by epigenetic and transcription factor (TF)-networks which are governed by bistable epigenetic switches [26]. Applying this model to cell differentiation data we recently identified situations where variant transcription of genes is accompanied by invariant epigenetic promoter states or vice versa. Interestingly, the former situation of TF-dominated regulation seems to apply to elementary cell functions related to stress response, cell cycle regulation and cell metabolism and requires mostly high expression levels of the genes beyond the sensitivity range of the switches. Combined regulation is found for developmental processes where genes become activated or deactivated by epigenetics, usually via histone methylation changes associated with DNA-hypo- or –hypermethylation near their promoters. Changes of methylation with only minor effect on transcription was found for *GPCRs* also upon cell differentiation. Overall we find striking agreement between gene functions in these three regimes between cell developmental data [26], WHO grade IV GBM [25] and the LGG studied here. Interestingly, methylation seems to-activate enhancers in TF-networks while it de-activates enhancers for developmental processes [70] or, in other words, enhancer and promoter methylation seem to act in an antagonistic fashion for both types of processes. A more simplistic interpretation of partly decoupled expression and methylation assumes rarely or non-overlapping sets of 'passenger' genes regulated by TFs and/or epigenetic 'drivers' such as the IDH-mutation [25].

On a cellular level, our results support a multi-component approach underpinned by single-cell transcriptome characteristics [56] that indicates variable composition of the tumors regarding healthy astrocyte- and oligodendrocyte-like cells, microglia/macrophage and proliferative stem cell-like constituents as illustrated in the right part of Figure 5A. The TME of the subtypes is characterized by marked variations of the immune cell composition that overlays with methylation changes of their genomes. It suggests an epigenetically-mediated interplay between tumor cells and immune cells in the TME.

We found footprints of previously published expression and methylation gene signatures extracted from studies on WHO grade II, III and IV gliomas in the tumors studied here indicating a considerable overlap of molecular mechanisms between LGG and GBM [20] in agreement with previous studies which underlined relevance of GBM molecular signatures for LGG [71]. These results support the view that the molecular heterogeneity of gliomas decomposes into a set of gene-regulatory modes that were activated in different combinations and to a different extent in the different subtypes and in tumors of different grades. In addition to the G-CIMP and G-CIMP-O signatures that typically occur in *IDH*-mut gliomas we also found methylation characteristics occurring in *IDH*-wt GBM such as the RTK-

II and mesenchymal methylation signatures reflecting concerted methylation changes of respective groups of genes in *IDH*-mut LGG as well (Figure 5A, part below). Moreover, we found concerted methylation patterns of the olfactory subgenome collecting *GPCR* genes and of cancer-related Keratin intermediate filament genes, respectively. These signatures overlap and combine in different ways giving rise to diverse methylation and expression patterns that partly shape the glioma phenotypes.

Phenotypic relatedness suggests developmental paths of gliomas

For a more detailed view on the relatedness between the subtypes we performed similarity tree analysis of the molecular tumor landscapes (Figure 5B). The expression and methylation 'phenotypic' trees obtained differ mainly in the position of the *IDH*-wt (C1) subtype. Its expression characteristics show rather similarities with the C2 tumors because of common inflammatory and astrocytic signatures while its methylation profiles rather resemble that of neuronal (C4) tumors owing to the common lack of the G-CIMP patterns (see the schematic profiles in the lower part of Figure 5B). On the other hand, both trees reflect similar mutual relations between the neuronal, oligodendroglioma-like (C3) and astrocytoma-like (C2) tumors where the former two types share similarities mainly regarding (low) inflammatory, (high) neuronal expression and (high) *GPCR*-methylation levels. Degeneration of apparent healthy brain functions in all subtypes, activated proliferation in C3 and partly inflammation in E3 seem to be driven by anti-correlated DNA-promoter methylation changes.

Interestingly, the astrocytoma-like subtypes in C2 sort in the order E4-E2-E3 and M4-M3-M2, respectively, which associates with increasing WHO grade of the tumors, their age at first diagnosis, their hazard ratio, the decrease of the global methylation levels and of neuronal expression characteristics and increased senescence bypass characteristics. We hypothesize that these trends reflect aspects of the progression of astrocytoma like gliomas from earlier to later phases in the natural course of the disease [20, 67]. Interestingly, these trends also suggest increasing therapy resistance along the proneural-to-mesenchymal axis after comparison with resistance and inflammatory signatures derived from GBM [22]. Search for glioma subtypes and/or molecular characteristics most suitable for immunotherapies is a challenge [59]. The inflammatory subtype E3 with maximum M2-macrophage polarization could be of interest for therapies targeting glioma associated macrophages [72].

Importantly, decreasing total methylation decomposes into reduced methylation of the *GPCR*- and Keratin methylation patterns on one hand, and the G-CIMP pattern, which shows the opposite trend in M4–M3, on the other hand (Figure 5B, part below). The relative reduction of G-CIMP in M2 is compatible with the observation that while IDH-mut associated G-CIMP initiates gliomagenesis it seems not required for later clonal expansions [73]. Interestingly, the RTKII-signature originally obtained from WHO grade IV *IDH*-wild type gliomas shows parallels with the senescence bypass signature in *IDH*-mutated LGG, and particularly reflects differences between Chr1p/19q-codeleted and –non-codeleted tumors. Both, loss and gain of methylation take place in parallel in different regions of the genome of tumor cells and/or in different cellular constituents of the TME. The subtype E5 manifests characteristics of early stages of astrocytoma-like tumors such as high levels of the *GPCR*- and Keratin-methylation patterns and low levels of the G-CIMP-methylation in addition to the expression of cancer testis genes. E5 collects both, Chr1p/19q-non-codeleted (mainly grade II) and, to a less amount, Chr1p/19q-codel tumors, which suggests that mechanisms affecting DNA methylation act partly independent of the Chr1p/19q-codel status.

Aberrant methylation shapes glioma phenotypes

Gliomas are consistently characterized by the loss of neuronal expression signature, especially in *IDH*-wt and *IDH*-mut astrocytoma-like, and to a less degree, also in oligodendroglioma-like tumors,

paralleled by decreasing expression and hyper-methylation of *PRC2*-targets with possible consequences for senescence bypass mechanisms. The latter properties are hallmarks of CIMP-like subtypes observed in colon cancer and lymphoma and in grade IV glioma [10, 38]. Previous studies have proposed a role for *PRC2* genes in protecting neurons against degeneration by repressing aberrant transcriptional programs [51]. Stratification of repressed chromatin states in fetal and adult brain revealed an antagonistic methylation pattern between oligodendroglioma-like (C3) and *IDH*-wt gliomas and an intermediate pattern in astrocytoma-like (C2) tumors which suggests deregulation of developmental cellular programs in *IDH*-wt; and of programs of differentiated tissue in C3 and partly also in C2. The former effect associates with the activation of inflammation and mesenchymal characteristics while the second one seems to activate proliferation and oxphos metabolism. Epigenetic activation of otherwise suppressed cellular programs seems to be essential for glioma development and diversification into subtypes.

Our methylation analysis uses integral methylation beta-values of upstream regions of each gene which are assumed to reflect mean promoter methylation levels. Similar methylation patterns were found in extended upstream regions and in the gene body, which suggests that DNA-methylation affects widespread genomic regions in a similar fashion. On the other hand, this integral methylation analysis eventually overlooks local and CpG-specific methylation effects with possible impact for transcriptional regulation. In this context, our integral method should be judged as one limitation of this study. We expect that alternative methods will further improve our understanding of the role of DNA-methylation, e.g., to better resolve the regulatory element landscapes and transcription factor networks [70] in gliomas. Also the possible impact of methylation of the olfactory subgenome on cell function and glioma development remains partly unclear and requires future work.

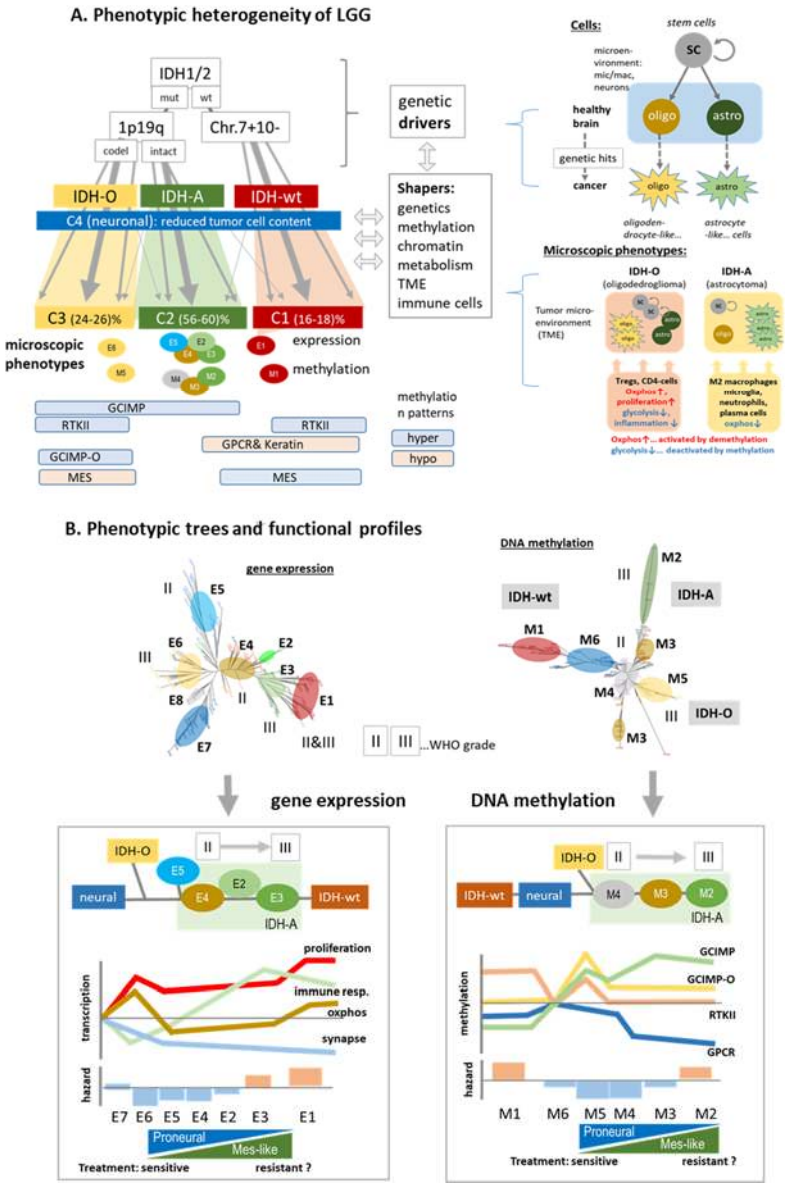


Figure 5: Schematic summary: A) The major glioma subtypes arise after specific genetic hits. The tumor phenotypes are then shaped by the tumor microenvironment (TME), its cell composition, epigenetics and additional genetic defects. Different methylation patterns develop in a subtype specific fashion upon tumor progression (left part). On a cellular level, astrocyte-like and oligodendrocyte-like gliomas are both primarily composed of proliferating stem cells, oligodendrocytes and astrocytes, however in different amounts, which associates with different immune cell compositions in the TME and metabolic expression signatures, which partly are affected by methylation effects. B) Phenotypic trees provide similarity relations between the expression and methylation subtypes (top), which were simplified as one-dimensional sequences of subtypes and associated with selected transcriptional programs, methylation patterns and prognosis (bottom).

Conclusions

Our study demonstrates the importance of molecular subtyping of LGG as a concept to better understand the biology of this disease. IDH mutated astrocytoma-like LGGs constitute the most heterogeneous sub-entity, which stratifies into distinct transcriptomic and methylation subtypes with possible impact for clinics, e.g. for identification of treatment resistant or sensitive tumor strata.

Analogies between astrocytoma-like LGG with grade IV *IDH*-wt tumors regarding varying treatment resistance suggest similar disease mechanisms; however further studies are required for verification. Hereby epigenetics, and particularly, DNA methylation is a shaping and driving factor of glioma heterogeneity and progression. Genomic, epigenetic and transcriptomic factors act in concert but partly also in a decoupled fashion what underpins the need for integrative, multidimensional subtyping of LGG by combining these data on gene and cellular levels in order to delineate mechanisms of gene (de-)regulation and to enable better patient stratification for individualization of treatment.

List of abbreviations

2HG	2-hydroxyglutarate
BP	biological process
C1 – C4	consensus classes of LGG with mutual correspondence between expression and methylation
CL	classical subtype of grade IV gliomas according to (Verhaak, et al., 2010)
CN	copy number
E1 – E8	expression subtypes of LGG
EMT	epithelial-mesenchymal transition
EPL	early-progenitor-like subtype
G-CIMP	Glioma CpG-Island hyperMethylation Phenotype
G-CIMP-O	G-CIMP with specific hypermethylation of *IDH*-O
G-CIMP-wt	hypermethylation of *IDH*-wt, particularly of the RTK II type
GEO	gene expression omnibus database
GGN	German Glioma Network
GPCR	G-protein coupled receptor
GSZ	gene set Z-score
HM	Hallmarks of Cancer
HR	hazard ratio
IDH	isocitrate dehydrogenase
IDH-A	malignant astrocytoma-like
IDH-mut-codel	*IDH* mutant additionally carrying codeletion of chromosome arm 1p and 19q
IDH-mut non-codel	*IDH* mutant with euploid 1p/19q
IDH-O	malignant oligodendrocytoma-like
IDH-wt	*IDH* wild-type
LGG	lower grade gliomas (WHO grade II and III)
M1 – M6	methylation subtypes
ME	mesenchymal subtype of grade IV gliomas according to (Verhaak, et al., 2010)
MGMT	O-6-methylguanine-DNA methyltransferase
MP	methylation pattern
MS	mesenchymal subtype (Sturm et al. (reanalyzed by Hopp et al.))
NL	neural subtype of grade IV gliomas according to (Verhaak, et al., 2010)
NPC	neural progenitor cells
OS HR	overall survival hazard ratio
Ox	oxidative phosphorylation
PG	pre-glioblastoma subtype of LGG (Gorovets, et al., 2012)
PN	proneural subtype of grade IV gliomas according to (Verhaak, et al., 2010)
PRC2	polycomb repressive complex 2
RTK	receptor tyrosine kinase
RTKI	GBM methylation class (Sturm et al., 2012)
RTKII	GBM methylation class (Sturm et al., 2012)
SOM	self-organizing map

TF	transcription factors
TME	tumor microenvironment

Declarations

Ethics approval and consent to participate

All patients gave written informed consent for participation in the GGN and its translational research projects.

Consent for publication

Not applicable.

Availability of data and materials

Expression data are available in the gene expression omnibus (GEO) database under accession number GSE61374 [6]. DNA methylation data reported in this study for the first time are deposited in the gene expression omnibus (GEO) database under accession number /to be added after acceptance/. The complete data are available from 'The Leipzig Health Atlas' repository (accession number will be available after acceptance of the manuscript).

Competing interests

MW has received research grants from Abbvie, Acceleron, Actelion, Bayer, Merck, Sharp & Dohme (MSD), Merck (EMD), Novocure, OGD2, Piqur, Roche and Tragara, and honoraria for lectures or advisory board participation or consulting from Abbvie, BMS, Celgene, Celldex, Merck, Sharp & Dohme (MSD), Merck (EMD), Novocure, Orbus, Pfizer, Progenics, Roche, Teva and Tocagen. US has received honoraria for lectures or advisory board participation from medac, GSK, mundipharma, Novartis, Novocure, Roche. DK has received research grants from Novocure, Northwest biotherapeutics, Kyowa, and honoraria for lectures or advisory board participation or consulting from Baxter and Kyowa. WW received study drug for clinical research from Apogenix, Roche and Pfizer. JCT has received research grants BrainLab and honoraria for lectures from BrainLab, Siemens, Merck, Roche and medac.

All other authors declare that they have no competing interests.

Funding

This work is supported by the Federal Ministry of Education and Research (BMBF), project grant SysGlio and the DKTK joint funding project "Next generation molecular diagnostics of malignant gliomas" and the German Glioma Network (GGN).

Authors' contributions

Coordinator of GGN: MW. Conceived and performed analysis, wrote manuscript: HB, EW. Performed analyses, contributed to the study and the manuscript: all authors. All authors read and approved the final manuscript.

Acknowledgments

Additional materials

Supplementary Table S 1: List of GGN-patients whose DNA methylation data were included into the study (separate Excel-file).

Supplementary Table S 2: List of TCGA-samples used for verification analyses (separate Excel-file).

Supplementary file 1: contains all supplementary figures and supplementary tables S3 – S6

References

1. Weller M, Wick W, Aldape K, Brada M, Berger M, Pfister SM, Nishikawa R, Rosenthal M, Wen PY, Stupp R, Reifenberger G: **Glioma.** In *Nature reviews Disease primers*, vol. 1. pp. 15017; 2015:15017.
2. Eckel-Passow JE, Lachance DH, Molinaro AM, Walsh KM, Decker PA, Sicotte H, Pekmezci M, Rice T, Kosel ML, Smirnov IV, et al: **Glioma Groups Based on 1p/19q, IDH, and TERT Promoter Mutations in Tumors.** *New England Journal of Medicine* 2015, **372:**2499-2508.
3. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW: **The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary.** *Acta Neuropathologica* 2016, **131:**803-820.
4. Suzuki H, Aoki K, Chiba K, Sato Y, Shiozawa Y, Shiraishi Y, Shimamura T, Niida A, Motomura K, Ohka F, et al: **Mutational landscape and clonal architecture in grade II and III gliomas.** *Nat Genet* 2015, **47:**458-468.
5. Network TCGAR: **Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas.** *New England Journal of Medicine* 2015, **372:**2481-2498.
6. Weller M, Weber R, Willscher E, Riehmer V, Hentschel B, Kreuz M, Felsberg J, Beyer U, Löffler-Wirth H, Kaulich K, et al: **Molecular classification of diffuse cerebral WHO grade II/III gliomas using genome- and transcriptome-wide profiling improves stratification of prognostically distinct patient groups.** *Acta Neuropathologica* 2015**:**1-15.
7. Reifenberger G, Weber RG, Riehmer V, Kaulich K, Willscher E, Wirth H, Gietzelt J, Hentschel B, Westphal M, Simon M, et al: **Molecular characterization of long-term survivors of glioblastoma using genome- and transcriptome-wide profiling.** *International Journal of Cancer* 2014, **135:**1822-1831.
8. Das PM, Singal R: **DNA Methylation and Cancer.** *Journal of Clinical Oncology* 2004, **22:**4632-4642.
9. Klutstein M, Nejman D, Greenfield R, Cedar H: **DNA Methylation in Cancer and Aging.** *Cancer Research* 2016, **76:**3446-3450.
10. Martinez R, Martin-Subero JI, Rohde V, Kirsch M, Alaminos M, Fernández AF, Ropero S, Schackert G, Esteller M: **A microarray-based DNA methylation study of glioblastoma multiforme.** *Epigenetics* 2009, **4:**255-264.
11. Laffaire J, Everhard S, Idbaih A, Crinière E, Marie Y, de Reyniès A, Schiappa R, Mokhtari K, Hoang-Xuan K, Sanson M, et al: **Methylation profiling identifies 2 groups of gliomas according to their tumorigenesis.** *Neuro-Oncology* 2011, **13:**84-98.
12. Kloosterhof NK, Bralten LBC, Dubbink HJ, French PJ, van den Bent MJ: **Isocitrate dehydrogenase-1 mutations: a fundamentally new understanding of diffuse glioma?** *The Lancet Oncology* 2011, **12:**83-91.
13. Christensen BC, Smith AA, Zheng S, Koestler DC, Houseman EA, Marsit CJ, Wiemels JL, Nelson HH, Karagas MR, Wrensch MR, et al: **DNA Methylation, Isocitrate Dehydrogenase Mutation, and Survival in Glioma.** *Journal of the National Cancer Institute* 2011, **103:**143-153.
14. Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP, Pan F, Pelloski CE, Sulman EP, Bhat KP, et al: **Identification of a CpG Island Methylator Phenotype that Defines a Distinct Subgroup of Glioma.** *Cancer Cell* 2010, **17:**510-522.

15. Brennan Cameron W, Verhaak Roel GW, McKenna A, Campos B, Noushmehr H, Salama Sofie R, Zheng S, Chakravarty D, Sanborn JZ, Berman Samuel H, et al: **The Somatic Genomic Landscape of Glioblastoma.** *Cell* 2013, **155:**462-477.
16. Sturm D, Witt H, Hovestadt V, Khuong-Quang D-A, Jones David TW, Konermann C, Pfaff E, Tönjes M, Sill M, Bender S, et al: **Hotspot Mutations in H3F3A and IDH1 Define Distinct Epigenetic and Biological Subgroups of Glioblastoma.** *Cancer Cell* 2012, **22:**425-437.
17. de Souza CF, Sabedot TS, Malta TM, Stetson L, Morozova O, Sokolov A, Laird PW, Wiznerowicz M, Iavarone A, Snyder J, et al: **A Distinct DNA Methylation Shift in a Subset of Glioma CpG Island Methylator Phenotypes during Tumor Recurrence.** *Cell Reports* 2018, **23:**637-651.
18. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, Koelsche C, Sahm F, Chavez L, Reuss DE, et al: **DNA methylation-based classification of central nervous system tumours.** *Nature* 2018, **555:**469.
19. Ward PS, Cross JR, Lu C, Weigert O, Abel-Wahab O, Levine RL, Weinstock DM, Sharp KA, Thompson CB: **Identification of additional IDH mutations associated with oncometabolite R(-)-2-hydroxyglutarate production.** *Oncogene* 2012, **31:**2491-2498.
20. Ceccarelli M, Barthel Floris P, Malta Tathiane M, Sabedot Thais S, Salama Sofie R, Murray Bradley A, Morozova O, Newton Y, Radenbaugh A, Pagnotta Stefano M, et al: **Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma.** *Cell* 2016, **164:**550-563.
21. Mazor T, Pankov A, Johnson BE, Hong C, Hamilton EG, Bell RJA, Smirnov IV, Reis GF, Phillips JJ, Barnes MJ, et al: **DNA methylation and somatic mutations converge on cell cycle and define similar evolutionary histories in brain tumors.** *Cancer Cell* 2015, **28:**307-317.
22. Segerman A, Niklasson M, Haglund C, Bergström T, Jarvius M, Xie Y, Westermark A, Sönmez D, Hermansson A, Kastemar M, et al: **Clonal Variation in Drug and Radiation Response among Glioma-Initiating Cells Is Linked to Proneural-Mesenchymal Transition.** *Cell Reports* 2016, **17:**2994-3009.
23. Klughammer J, Kiesel B, Roetzer T, Fortelny N, Nemc A, Nenning K-H, Furtner J, Sheffield NC, Datlinger P, Peter N, et al: **The DNA methylation landscape of glioblastoma disease progression shows extensive heterogeneity in time and space.** *Nature Medicine* 2018.
24. Wang Q, Hu B, Hu X, Kim H, Squatrito M, Scarpace L, deCarvalho AC, Lyu S, Li P, Li Y, et al: **Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment.** *Cancer Cell* 2017, **32:**42-56.e46.
25. Hopp L, Löffler-Wirth H, Galle J, Binder H: **Combined SOM-portrayal of gene expression and DNA methylation landscapes disentangles modes of epigenetic regulation in glioblastoma.** *Epigenomics* 2018, **10:**745-764.
26. Thalheim T, Hopp L, Binder H, Aust G, Galle J: **On the Cooperation between Epigenetics and Transcription Factor Networks in the Specification of Tissue Stem Cells.** *Epigenomes* 2018, **2:**20.
27. Wirth H, Löffler M, von Bergen M, Binder H: **Expression cartography of human tissues using self organizing maps.** *BMC Bioinformatics* 2011, **12:**306.
28. Löffler-Wirth H, Kalcher M, Binder H: **oposSOM: R-package for high-dimensional portraying of genome-wide expression landscapes on bioconductor.** *Bioinformatics* 2015, **31:**3225-3227.
29. Kunz M, Löffler-Wirth H, Dannemann M, Willscher E, Doose G, Kelso J, Kottek T, Nickel B, Hopp L, Landsberg J, et al: **RNA-seq analysis identifies different transcriptomic types and developmental trajectories of primary melanomas.** *Oncogene* 2018.
30. Toronen P, Ojala P, Marttinen P, Holm L: **Robust extraction of functional signals from gene set analysis using a generalized threshold free scoring function.** *BMC Bioinformatics* 2009, **10:**307.
31. Wirth H, von Bergen M, Binder H: **Mining SOM expression portraits: Feature selection and integrating concepts of molecular function** *BioData Mining* 2012, **5:18**.

32. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA: **Robust enumeration of cell subsets from tissue expression profiles.** *Nature Methods* 2015, **12:**453.
33. Brat DJ, Aldape K, Colman H, Holland EC, Louis DN, Jenkins RB, Kleinschmidt-DeMasters BK, Perry A, Reifenberger G, Stupp R, et al: **cIMPACT-NOW update 3: recommended diagnostic criteria for "Diffuse astrocytic glioma, IDH-wildtype, with molecular features of glioblastoma, WHO grade IV".** *Acta Neuropathologica* 2018, **136:**805-810.
34. Glusman G, Yanai I, Rubin I, Lancet D: **The Complete Human Olfactory Subgenome.** *Genome Research* 2001, **11:**685-702.
35. Verhaak RGW, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, et al: **Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1.** *Cancer Cell* 2010, **17:**98-110.
36. Gorovets D, Kannan K, Shen R, Kastenhuber ER, Islamdoust N, Campos C, Pentsova E, Heguy A, Jhanwar SC, Mellinghoff IK, et al: **IDH Mutation and Neuroglial Developmental Features Define Clinically Distinct Subclasses of Lower Grade Diffuse Astrocytic Glioma.** *Clinical Cancer Research* 2012, **18:**2490-2501.
37. Paul Y, Mondal B, Patil V, Somasundaram K: **DNA methylation signatures for 2016 WHO classification subtypes of diffuse gliomas.** *Clinical epigenetics* 2017, **9:**32-32.
38. Hopp L, Willscher E, Wirth-Loeffler H, Binder H: **Function Shapes Content: DNA-Methylation Marker Genes and their Impact for Molecular Mechanisms of Glioma.** *Journal of Cancer Research Updates* 2015, **4:**127-148.
39. Yao C, Li H, Shen X, He Z, He L, Guo Z: **Reproducibility and Concordance of Differential DNA Methylation and Gene Expression in Cancer.** *PLOS one* 2012, **7:**e29686.
40. Huang K, Shen Y, Xue Z, Bibikova M, April C, Liu Z, Cheng L, Nagy A, Pellegrini M, Fan J-B, Fan G: **A Panel of CpG Methylation Sites Distinguishes Human Embryonic Stem Cells and Induced Pluripotent Stem Cells.** *Stem Cell Reports* 2014, **2:**36-43.
41. Polioudaki H, Agelaki S, Chiotaki R, Politaki E, Mavroudis D, Matikas A, Georgoulias V, Theodoropoulos PA: **Variable expression levels of keratin and vimentin reveal differential EMT status of circulating tumor cells and correlation with clinical characteristics and outcome of patients with metastatic breast cancer.** *BMC Cancer* 2015, **15:**399.
42. Karantza V: **Keratins in health and cancer: more than mere epithelial cell markers.** *Oncogene* 2011, **30:**127-138.
43. Hofmann O, Caballero OL, Stevenson BJ, Chen Y-T, Cohen T, Chua R, Maher CA, Panji S, Schaefer U, Kruger A, et al: **Genome-wide analysis of cancer/testis gene expression.** *Proceedings Of The National Academy Of Sciences Of The United States Of America* 2008, **105:**20422-20427.
44. Ghafouri-Fard S, Modarressi M-H: **Expression of cancer–testis genes in brain tumors: implications for cancer immunotherapy.** *Immunotherapy* 2012, **4:**59-75.
45. Freitas MRP, Malheiros SMF, Stávale JN, Biassi TP, Zamunér FT, de Souza Begnami MDF, Soares FA, Vettore AL: **Expression of Cancer/Testis Antigens is Correlated with Improved Survival in Glioblastoma.** *Oncotarget* 2013, **4:**636-646.
46. Yawata T, Nakai E, Park KC, Chihara T, Kumazawa A, Toyonaga S, Masahira T, Nakabayashi H, Kaji T, Shimizu K: **Enhanced expression of cancer testis antigen genes in glioma stem cells.** *Molecular Carcinogenesis* 2010, **49:**532-544.
47. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim T-K, Koche RP, et al: **Genome-wide maps of chromatin state in pluripotent and lineage-committed cells.** *Nature* 2007, **448:**553-560.
48. Hebenstreit D, Fang M, Gu M, Charoensawan V, van Oudenaarden A, Teichmann SA: **RNA sequencing reveals two major classes of gene expression levels in metazoan cells.** *Mol Syst Biol* 2011, **7**.

49. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, Weinberg RA: **An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors.** *Nat Genet* 2008, **40:**499-507.
50. Meissner A: **Epigenetic modifications in pluripotent and differentiated cells.** *Nat Biotech* 2010, **28:**1079-1088.
51. von Schimmelmann M, Feinberg PA, Sullivan JM, Ku SM, Badimon A, Duff MK, Wang Z, Lachmann A, Dewell S, Ma'ayan A, et al: **Polycomb repressive complex 2 (PRC2) silences genes responsible for neurodegeneration.** *Nature Neuroscience* 2016, **19:**1321.
52. Yang XW: **Life and death rest on a bivalent chromatin state.** *Nature Neuroscience* 2016, **19:**1271.
53. Ernst J, Kellis M: **Discovery and characterization of chromatin states for systematic annotation of the human genome.** *Nat Biotech* 2010, **28:**817-825.
54. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, et al: **Integrative analysis of 111 reference human epigenomes.** *Nature* 2015, **518:**317-330.
55. Hambardzumyan D, Gutmann DH, Kettenmann H: **The role of microglia and macrophages in glioma maintenance and progression.** *Nat Neurosci* 2016, **19:**20-27.
56. Venteicher AS, Tirosh I, Hebert C, Yizhak K, Neftel C, Filbin MG, Hovestadt V, Escalante LE, Shaw ML, Rodman C, et al: **Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq.** *Science* 2017, **355**.
57. Roesch S, Rapp C, Dettling S, Herold-Mende C: **When Immune Cells Turn Bad—Tumor-Associated Microglia/Macrophages in Glioma.** *International Journal of Molecular Sciences* 2018, **19:**436.
58. Domingues P, González-Tablas M, Otero Á, Pascual D, Miranda D, Ruiz L, Sousa P, Ciudad J, Gonçalves JM, Lopes MC, et al: **Tumor infiltrating immune cells in gliomas and meningiomas.** *Brain, Behavior, and Immunity* 2016, **53:**1-15.
59. Morisse MC, Jouannet S, Dominguez-Villar M, Sanson M, Idbaih A: **Interactions between tumor-associated macrophages and tumor cells in glioblastoma: unraveling promising targeted therapies.** *Expert Review of Neurotherapeutics* 2018, **18:**729-737.
60. Gabrusiewicz K, Rodriguez B, Wei J, Hashimoto Y, Healy LM, Maiti SN, Thomas G, Zhou S, Wang Q, Elakkad A, et al: **Glioblastoma-infiltrated innate immune cells resemble M0 macrophage phenotype.** *JCI Insight* 2016, **1**.
61. Huysentruyt LC, Akgoc Z, Seyfried TN: **Hypothesis: are neoplastic macrophages/microglia present in glioblastoma multiforme?** *ASN neuro* 2011, **3:**e00064.
62. Han S, Zhang C, Li Q, Dong J, Liu Y, Huang Y, Jiang T, Wu A: **Tumour-infiltrating CD4+ and CD8+ lymphocytes as predictors of clinical outcome in glioma.** *British Journal Of Cancer* 2014, **110:**2560.
63. Mu L, Yang C, Gao Q, Long Y, Ge H, DeLeon G, Jin L, Chang Y, Sayour EJ, Ji J, et al: **CD4+ and Perivascular Foxp3+ T Cells in Glioma Correlate with Angiogenesis and Tumor Progression.** *Frontiers in Immunology* 2017, **8:**1451.
64. Attarha S, Roy A, Westermark B, Tchougounova E: **Mast cells modulate proliferation, migration and stemness of glioma cells through downregulation of GSK3β expression and inhibition of STAT3 activation.** *Cellular Signalling* 2017, **37:**81-92.
65. Ooi YC, Tran P, Ung N, Thill K, Trang A, Fong BM, Nagasawa DT, Lim M, Yang I: **The role of regulatory T-cells in glioma immunology.** *Clinical Neurology and Neurosurgery* 2014, **119:**125-132.
66. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC: **Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer.** *Immunity* 2013, **39:**782 - 795.
67. Barthel FP, Wesseling P, Verhaak RGW: **Reconstructing the molecular life history of gliomas.** *Acta Neuropathologica* 2018, **135:**649-670.

68. Xie W, Kagiampakis I, Pan L, Zhang YW, Murphy L, Tao Y, Kong X, Kang B, Xia L, Carvalho FLF, et al: **DNA Methylation Patterns Separate Senescence from Transformation Potential and Indicate Cancer Risk.** *Cancer Cell* 2018, **33:**309-321.e305.
69. Watanabe T, Nobusawa S, Kleihues P, Ohgaki H: **IDH1 Mutations Are Early Events in the Development of Astrocytomas and Oligodendrogliomas.** *The American Journal of Pathology* 2009, **174:**1149-1153.
70. Yao L, Shen H, Laird PW, Farnham PJ, Berman BP: **Inferring regulatory element landscapes and transcription factor networks from cancer methylomes.** *Genome biology* 2015, **16:**105-105.
71. Guan X, Vengoechea J, Zheng S, Sloan AE, Chen Y, Brat DJ, O'Neill BP, de Groot J, Yust-Katz S, Yung W-KA, et al: **Molecular Subtypes of Glioblastoma Are Relevant to Lower Grade Glioma.** *PLOS ONE* 2014, **9:**e91216.
72. Cinzia Dello R, Lucia L, Lucio T, Pierluigi N, Grazia G, Colin KC: **Exploiting Microglial Functions for the Treatment of Glioblastoma.** *Current Cancer Drug Targets* 2017, **17:**267-281.
73. Mazor T, Chesnelong C, Pankov A, Jalbert LE, Hong C, Hayes J, Smirnov IV, Marshall R, Souza CF, Shen Y, et al: **Clonal expansion and epigenetic reprogramming following deletion or amplification of mutant IDH1.** *Proceedings of the National Academy of Sciences* 2017, **114:**10743-10748.
74. Gerber T, Willscher E, Loeffler-Wirth H, Hopp L, Schadendorf D, Schartl M, Anderegg U, Camp G, Treutlein B, Binder H, Kunz M: **Mapping heterogeneity in patient-derived melanoma cultures by single-cell RNA-seq.** *Oncotarget* 2016, **8:**846-862.
75. Lessin SR, Huebner K, Isobe M, Croce CM, Steinert PM: **Chromosomal Mapping of Human Keratin Genes: Evidence of Non-linkage.** *Journal of Investigative Dermatology* 1988, **91:**572-578.
76. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov Jill P, Tamayo P: **The Molecular Signatures Database Hallmark Gene Set Collection.** *Cell Systems* 2015, **1:**417-425.
77. Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE, Matese JC, Perou CM, Hurt MM, Brown PO, Botstein D: **Identification of Genes Periodically Expressed in the Human Cell Cycle and Their Expression in Tumors.** *Molecular Biology of the Cell* 2002, **13:**1977-2000.
78. Hopp L, Wirth H, Fasold M, Binder H: **Portraying the expression landscapes of cancer subtypes: A glioblastoma multiforme and prostate cancer case study.** *Systems Biomedicine* 2013, **1:**99-121.
79. Donson AM, Birks DK, Schittone SA, Kleinschmidt-DeMasters BK, Sun DY, Hemenway MF, Handler MH, Waziri AE, Wang M, Foreman NK: **Increased Immune Gene Expression and Immune Cell Infiltration in High-Grade Astrocytoma Distinguish Long-Term from Short-Term Survivors.** *The Journal of Immunology* 2012, **189:**1920-1927.
80. Dong H, Siu H, Luo L, Fang X, Jin L, Xiong M: **Investigation gene and microRNA expression in glioblastoma.** *BMC Genomics* 2010, **11:**S16.
81. Mukasa A, Ueki K, Ge X, Ishikawa S, Ide T, Fujimaki T, Nishikawa R, Asai A, Kirino T, Aburatani H: **Selective Expression of a Subset of Neuronal Genes in Oligodendroglioma with Chromosome 1p Loss.** *Brain Pathology* 2004, **14:**34-42.
82. Shinawi T, Hill VK, Krex D, Schackert G, Gentle D, Morris MR, Wei W, Cruickshank G, Maher ER, Latif F: **DNA methylation profiles of long- and short-term glioblastoma survivors.** *Epigenetics* 2013, **8:**149-156.

Supplementary Text

DNA methylation, transcriptome and genetic copy number signatures of diffuse cerebral WHO grade II/III gliomas resolve cancer heterogeneity and development

H. Binder[#§1], E. Willscher[§1], H. Loeffler-Wirth[1], L. Hopp[1], D. T. W. Jones[2,3], S. M. Pfister[2,4,5], M. Kreuz[6], D. Gramatzki[7], E. Fortenbacher[1], B. Hentschel[6], M. Tatagiba[8], U. Herrlinger[9], H. Vatter[9], J. Matschke[10], M. Westphal[11], D. Krex[12], G. Schackert[12], J.C. Tonn[13], U. Schlegel[14], H.-J. Steiger[15], W. Wick[16,17], R. G. Weber[18], M. Weller[+7], M. Loeffler[+6]

1 Interdisciplinary Centre for Bioinformatics, Universität Leipzig, Härtelstr. 16–18, 04107 Leipzig, Germany

2 Hopp Children's Cancer Center Heidelberg (KiTZ), Im Neuenheimer Feld 430, 69120 Heidelberg, Germany

3 Pediatric Glioma Research Group, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

4 Division of Pediatric Neurooncology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

5 Department of Pediatric Oncology, Hematology and Immunology, Heidelberg University Hospital, Im Neuenheimer Feld 430, 69120 Heidelberg, Germany

6 Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Härtelstraße 16-18, 04107 Leipzig, Germany

7 Department of Neurology, University Hospital and University Zurich, Frauenklinikstrasse 26, 8091 Zurich, Switzerland

8 Clinic for Neurosurgery, Tübingen University Hospital, Hoppe-Seyler-Straße 3, 72076 Tübingen, Germany

9 Division of Clinical Neurooncology, Department of Neurology, University Hospital Bonn, Bonn, Germany

10 Institute of Neuropathologie, University Clinic Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany

11 Department of Neurosurgery, University Clinic Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany

12 Department of Neurosurgery, Technical University Dresden, Fetscherstraße 74, 01307 Dresden, Germany

13 Department of Neurosurgery, Ludwig Maximilians University Munich and German Cancer Consortium (DKTK), partner site Munich, Marchioninistraße 15, D-81377 Munich, Germany

14 Department of Neurology, University Hospital Knappschaftskrankenhaus Bochum-Langendreer, In der Schornau 23-25, 44892 Bochum, Germany

15 Clinic for Neurosurgery, University Düsseldorf, Moorenstr. 5, 40225 Düsseldorf, Germany

16 Clinical Cooperation Unit Neurooncology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

17 Neurology Clinic and National Center for Tumor Diseases, University Hospital Heidelberg, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany

18 Department of Human Genetics, Hannover Medical School, Carl-Neuberg-Str. 1, 30625 Hannover, Germany

Supplementary Figures

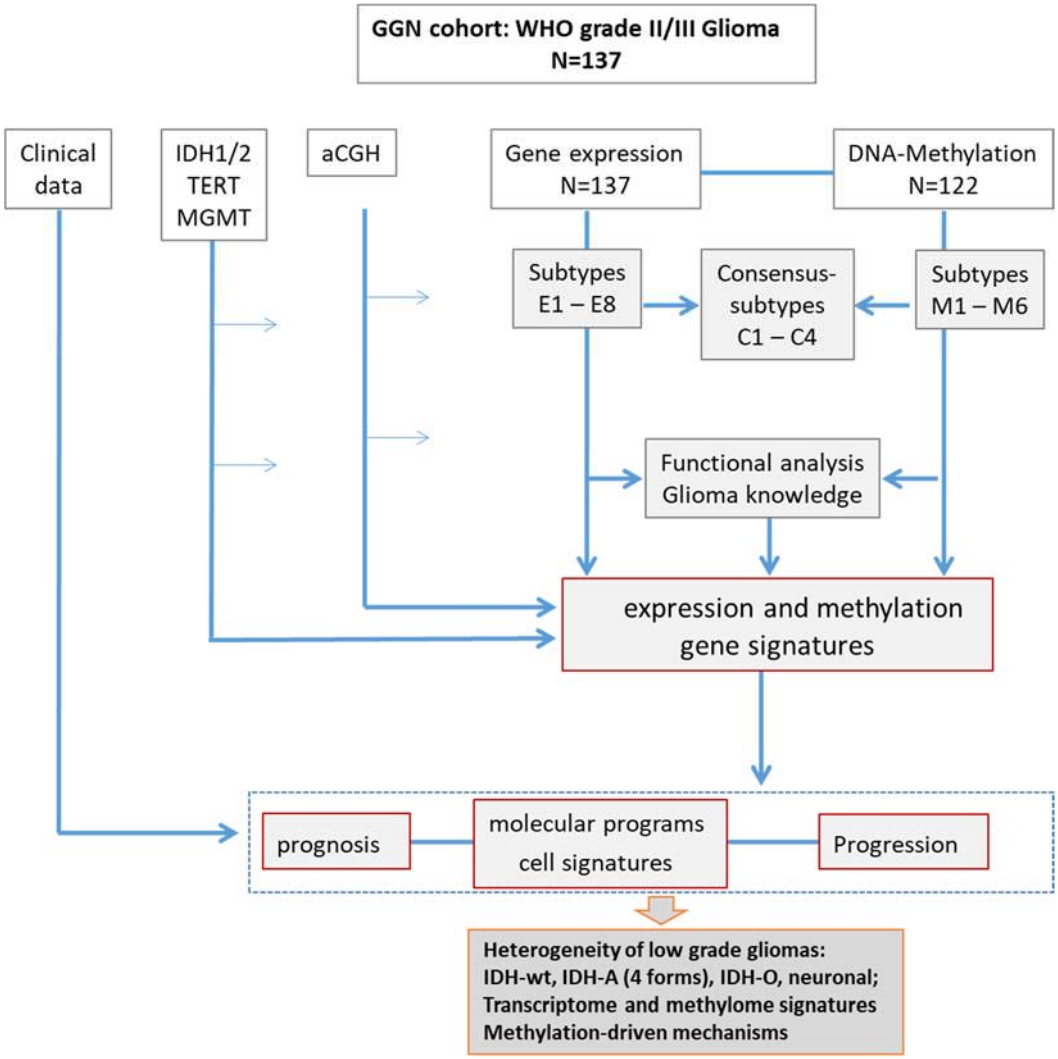


Figure S 1: Overview scheme summarizing data used, analyses and major objectives of the study

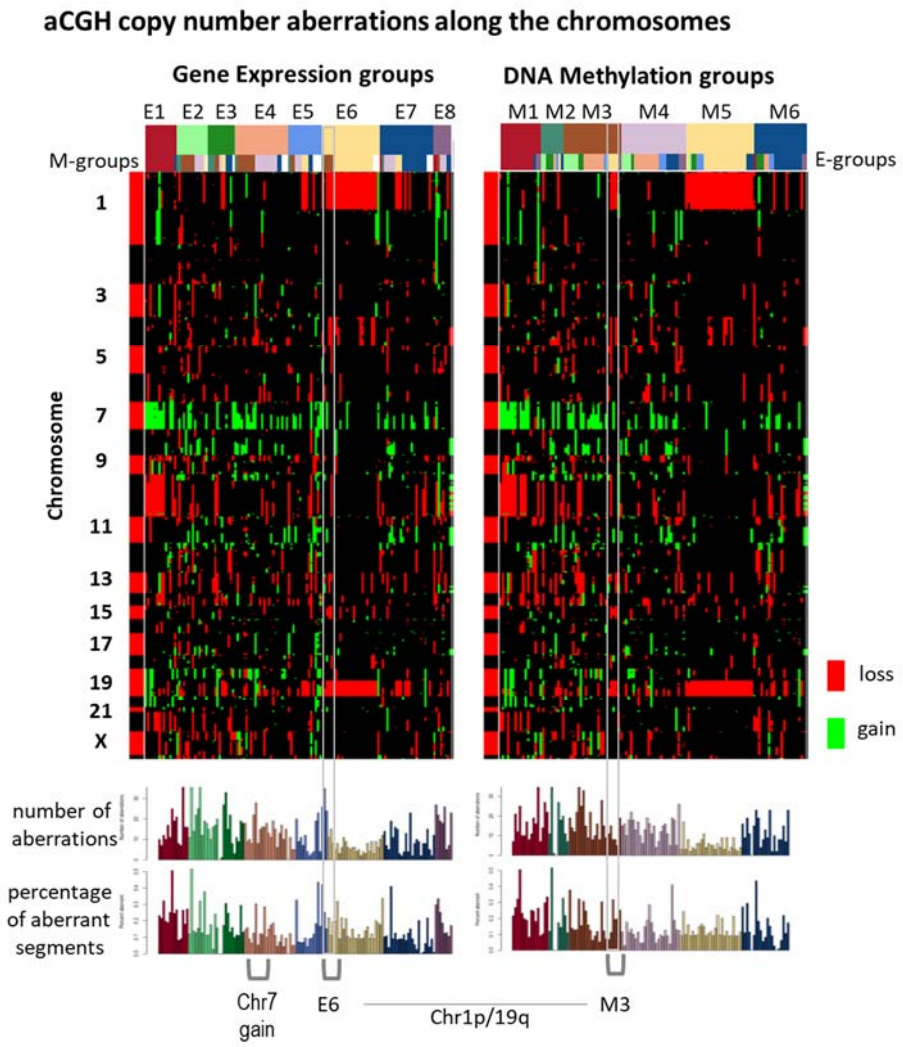


Figure S 2: Heatmap of copy number alterations in E- and M-groups of glioma samples. Copy number gains on Chr7 and losses on Chr10 accumulate in groups E1 and M1, whereas codeletions on Chr1p and Chr19q are frequently found in groups E6 and M5 which also are characterized by relative low total numbers of alterations. Copy number gains on Chr7 without Chr10 loss accumulate in samples of E4 and M3. Another combination is observed for E6 and M5 collecting samples with codeletions on Chr1p and Chr19q. Part of samples with these codeletions from E6 however sort into M3 (instead of M5, see the white frames in the figure) due to the different methylation of the olfactory subgenome (*GPCR*-signature) as indicated by the respective frames in Figure S 3.

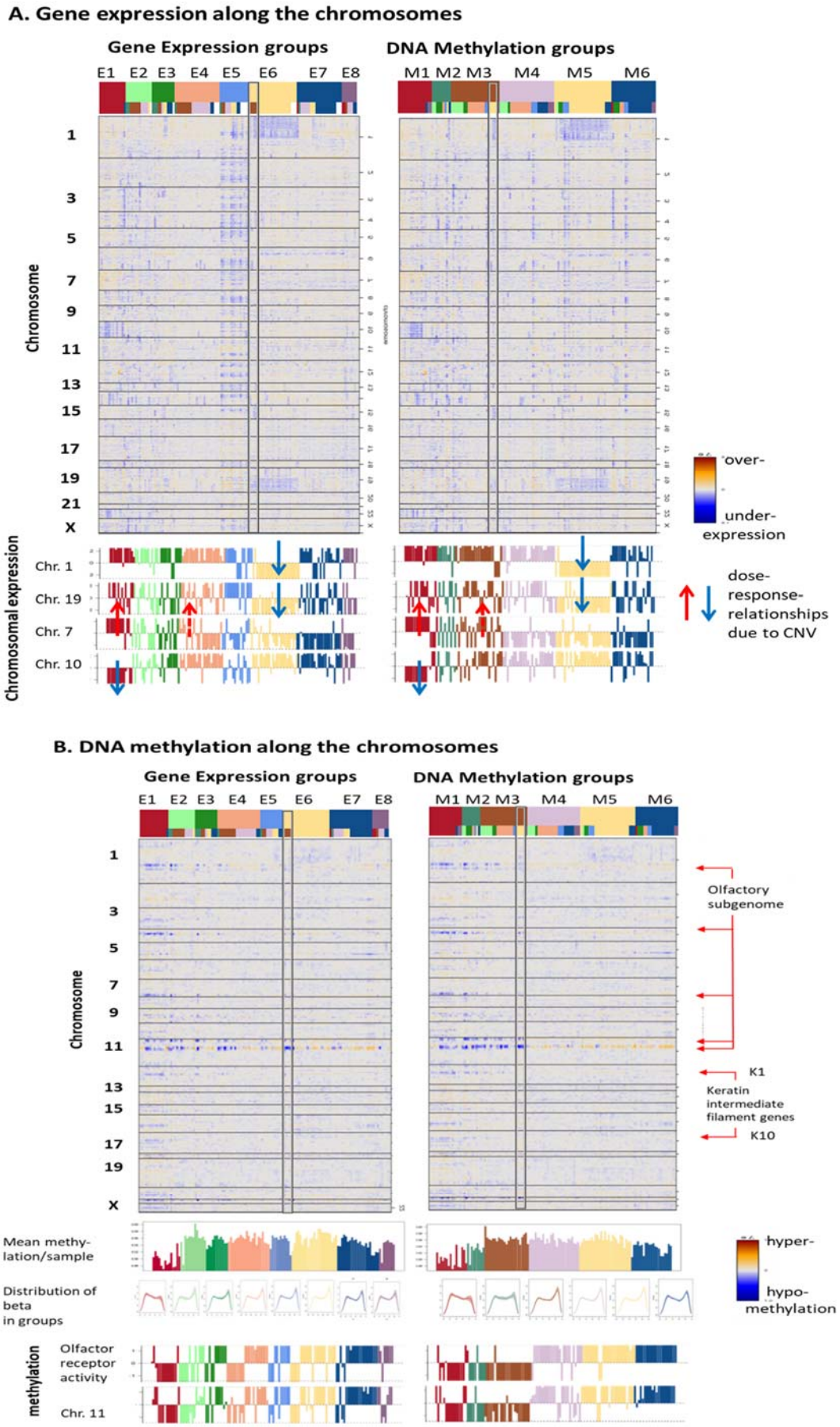

A. Gene expression along the chromosomes
Gene Expression groups
DNA Methylation groups
E1 E2 E3 E4 E5 E6 E7 E8
M1 M2 M3 M4 M5 M6
Chromosome
1
3
5
7
9
11
13
15
17
19
21
X
over-
under-
expression
Chromosomal expression
Chr. 1
Chr. 19
Chr. 7
Chr. 10
dose-
response-
relationships
due to CNV
B. DNA methylation along the chromosomes
Gene Expression groups
DNA Methylation groups
E1 E2 E3 E4 E5 E6 E7 E8
M1 M2 M3 M4 M5 M6
Chromosome
1
3
5
7
9
11
13
15
17
19
X
Olfactory
subgenome
K1
Keratin
intermediate
filament genes
K10
Mean methy-
lation/sample
Distribution of
beta
in groups
hyper-
hypo-
methylation
methylation
Olfactor
receptor
activity
Chr. 11

Figure S 3: Gene expression (part A) and methylation (part B) along the chromosomes was estimated using sliding-windows averaging [74] to study positioning effects of the genes. Chromosomal gene expression (part A) reveals regions of increased and decreased transcription due to copy number aberrations especially on Chr1, 7, 10 and 19. The mean expression of all genes from the respective chromosomes clearly indicates up- and downregulation for copy number gains and losses, respectively (part A below, see arrows). Positioning effects of methylation (part B) are evident in regions of the olfactory subgenome especially on Chr11 [34] and of the keratin intermediate filament gene clusters K1 and K10 [75]. The methylation profile of the gene set 'olfactory receptor activity' containing 175 genes coding olfactory receptors (75 of them on Chr11) closely agrees with the methylation profile of all genes from Chr. 11. The methylation profile of olfactory receptors reflects a sharp cut between methylation groups M1 to M3 (hypomethylated) and M4 to M6 (hypermethylated). The mean absolute methylation per sample is high for *IDH* mutated cases which accumulate in groups E2 to E5 and M2 to M5. The frequency distribution of beta values in general shows a bimodal shape with its maximum either at low or high beta for *IDH* wild type and mutated cases, respectively.

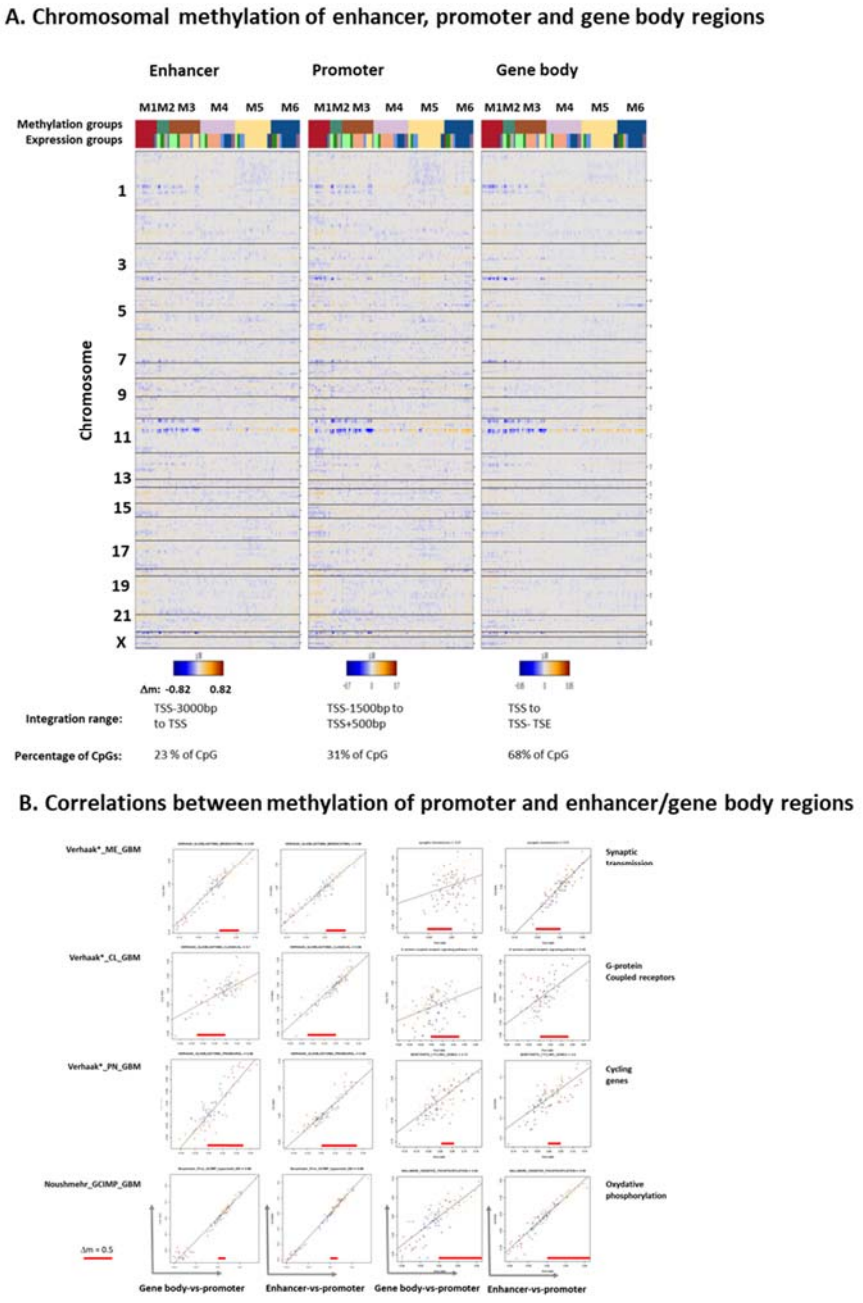


Figure S 4: The effect of DNA-methylation in gene enhancer, promoter and gene body regions: A) Methylation along the chromosomes is virtually identical on this rough scale. Particularly one finds hypomethylation of the olfactory subgenome in all three regions upstream and downstream of the TSS. B) Correlation plot between gene body/enhancer versus promoter regions for genes of different signature sets. For most of the signatures we find strong correlations (r^2>0.8) indicating similar mean methylation trends in enhancer, promoter and gene body regions. DNA methylation in the three different regions was calculated as the mean m-value (m= log [β/(1-β)]) averaged over CpG-probes over distinct genomic regions as indicated in part A of the figure (TSS…transcription start site; TSE…transcription end site). Overall these data suggest aberrant methylation in widespread regions of the genome. This notion is further supported by the result that marker CpGs randomly distribute over the three enhancer, promoter and gene body ranges in proportions 2:3:7 (Figure S 11).

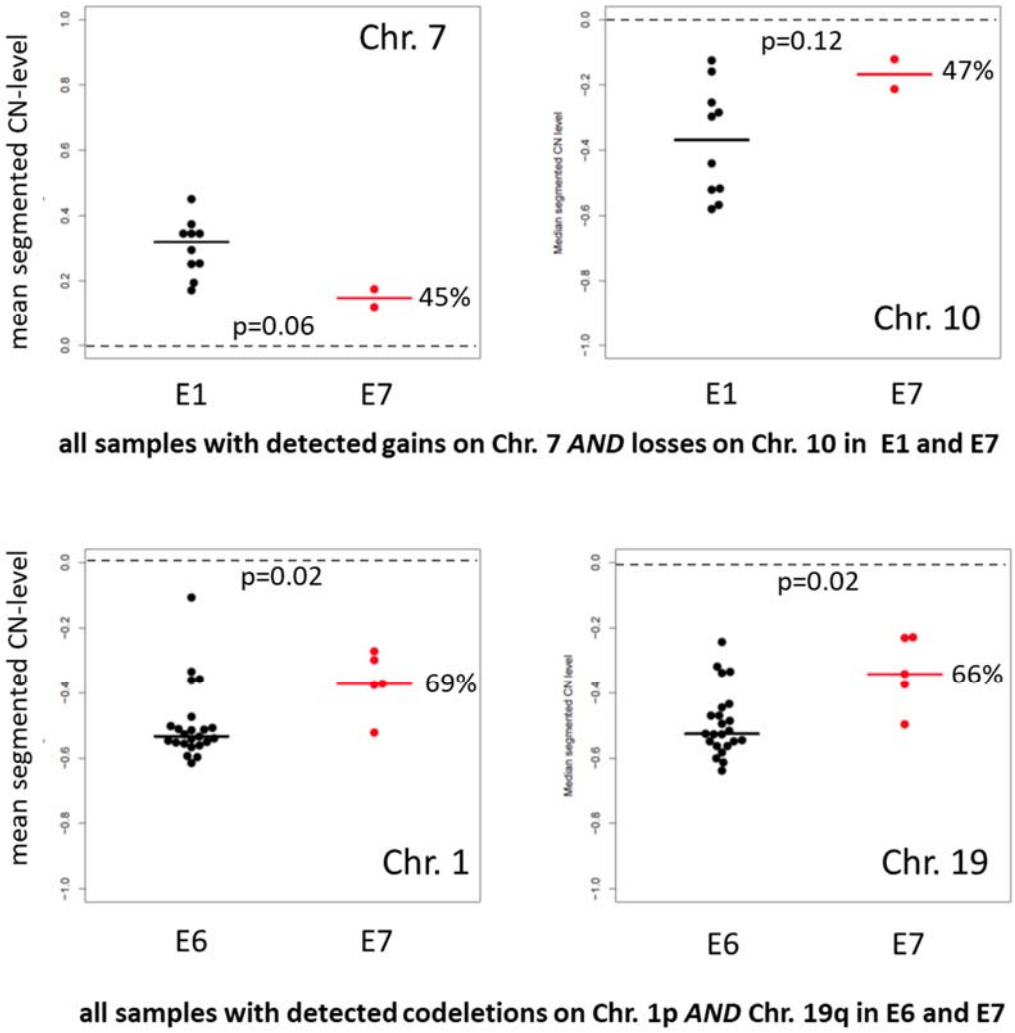
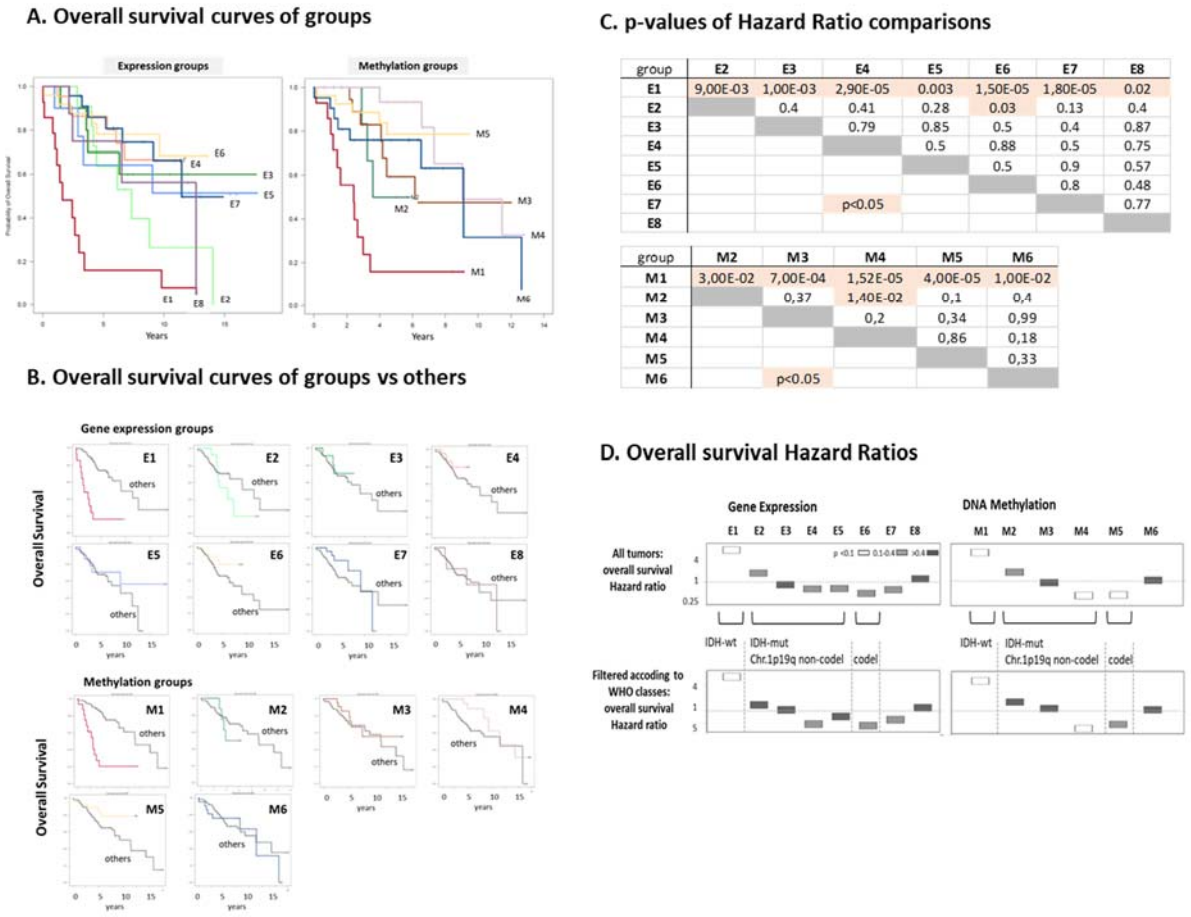


Figure S 5: Comparison of copy number (CN) levels on selected chromosomes between samples in subtype E7 and E1 showing gains and losses on Chr7 *and* Chr10, respectively, and between samples in subtypes E7 and E6 showing codeletions on Chr1p *and* Chr19q. The systematically smaller absolute levels of the CN in E7 compared with E1 and E6 suggests contaminations with healthy cells without these CN defects. Note that subtype E1 is strongly enriched in samples showing gains on Chr7 and losses on Chr10 (10 out of 14 samples) whereas subtype E6 strongly enriches in samples showing codeletions on Chr1p and Chr19q (24 out of 26 samples). On the other hand, subtype E7 (24 tumors in total) is a relatively heterogeneous mixture that contains samples showing either the one (2 samples with Chr7 gains and Chr10 losses) or the other (6 samples with Chr1p and Chr19q codeletions) key defect. The relative decrease of CN varies between (45-47)% (Chr7 gains and Chr10 losses) and (66-69)% (Chr1p and Chr19q codeletions) which roughly estimates the relative decrease of tumor cell content. p-values (t-test) for group differences were given in the figure. Note that expression analysis assigns gene signatures of healthy brain to E7 but not to E1 and E6 which supports the results of copy number analysis. Overall, these results support the view that E7 collects samples contaminated with healthy brain tissue.



group	E2	E3	E4	E5	E6	E7	E8
E1	9,00E-03	1,00E-03	2,90E-05	0.003	1,50E-05	1,80E-05	0.02
E2		0.4	0.41	0.28	0.03	0.13	0.4
E3			0.79	0.85	0.5	0.4	0.87
E4				0.5	0.88	0.5	0.75
E5					0.5	0.9	0.57
E6						0.8	0.48
E7			p<0.05				0.77
E8							

group	M2	M3	M4	M5	M6
M1	3,00E-02	7,00E-04	1,52E-05	4,00E-05	1,00E-02
M2		0,37	1,40E-02	0,1	0,4
M3			0,2	0,34	0,99
M4				0,86	0,18
M5					0,33
M6		p<0.05			

Figure S 6: Survival analysis of the E- and M-groups: A) Overall survival (OS) curves for E- and M-groups. Interestingly, the OS-curves of M-groups M3 – M6 seem to reflect more prognostic differences than the E-groups E3 – E8 which is supported also by the respective p-values (part C of the figure). However larger sample sizes are required for justification of this result. B) OS compared with that of the remaining samples of the cohort ('others'). C) Signifinace p-values of pairwise comparisons of hazard ratios (HR). E1 and M1 (IDH-wt) compared with the other groups (mostly IDH-mut) show significantly worse prognosis. E2 reveals unfavorable prognosis compared with the other E-groups. Similar prognosis applies to M2 among the M-groups. D) HRs of the groups were shown as boxplots. The HR data were calculated considering all tumors in each of the groups (upper row of figures) and after filtering the groups in accordance with WHO classifications (i.e., considering only IDH-wt in E1 and M1; IDH-mut and Chr.1p/19q non-codel in E2 – E5 and M2 – M4; IDH-mut and Chr.1p/19q codel in E6 and M5). The effect of filtering is small and doesn't change the general trends discussed in the paper.

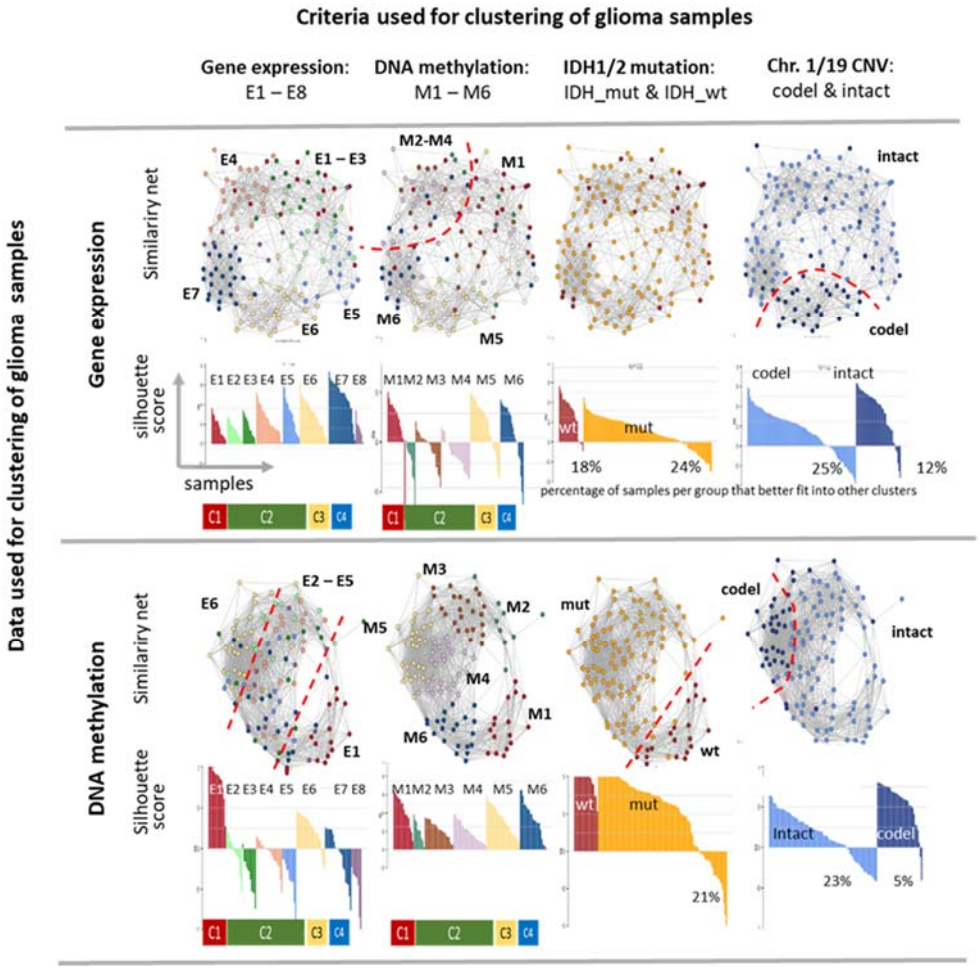


Figure S 7: Coloring of the samples according to different classification categories (see column headers) in the similarity network representation of the sample landscapes of expression (row above) and methylation (row below) data reveals different degrees of fuzziness of the class assignments: 1p/19q codeleted (and *IDH* mutated) tumors form clearly separate clusters in the methylation and expression landscapes as well. The *IDH* mutated and non-mutated tumors well separate each from another in the methylation landscape while their mutual distributions in the expression landscape are fuzzier. The samples of the expression subtypes E1 – E8 and that of the methylation subtypes M1 – M6 collect into almost severed clusters in the E- and M-nets, respectively. Here the assignment of samples to the subtypes meets the criterion of maximum silhouette scores used for classification. This score estimates the preference for each sample being member of the actual subtype as best choice compared with the second best choice as a member of another subtype using a correlation coefficient metrics [38]. The silhouette score estimates the optimal cluster membership of the samples with positive values indicating preference for the actual cluster chosen and negative values for alternative cluster memberships. It is always positive for the E- and M-subtypes using the expression and methylation data, respectively, thus indicating optimal clustering. Reversal of the labels, i.e. coloring of the samples in the E-net according to their M-label and vice versa shows still separate clouds for the samples subsumed in consensus classes C1, C3 and partly C4 but considerable mixing of the subtypes in C2 in agreement with the results presented in the main paper (Figure 1). The silhouette plots also show that typically about 20% of the samples with *IDH* mutations show closer similarities with the respective alternative clusters of non- *IDH* mutated tumors according to their expression and methylation patterns. Although the status of codeletions on Chr.1p/19q well separates the tumors in the E- and M-net as well about 5 – 25% of the samples of each group show closer similarity to the alternative group as indicated by the negative values of their silhouette score. Overall, this analysis shows that our subtyping shows clear associations between the gene expression and methylation landscapes and with the main genetic hallmarks. It is however overlaid by a certain degree of uncertainty reflecting intermixing of genetic, DNA methylation and transcriptional effects.

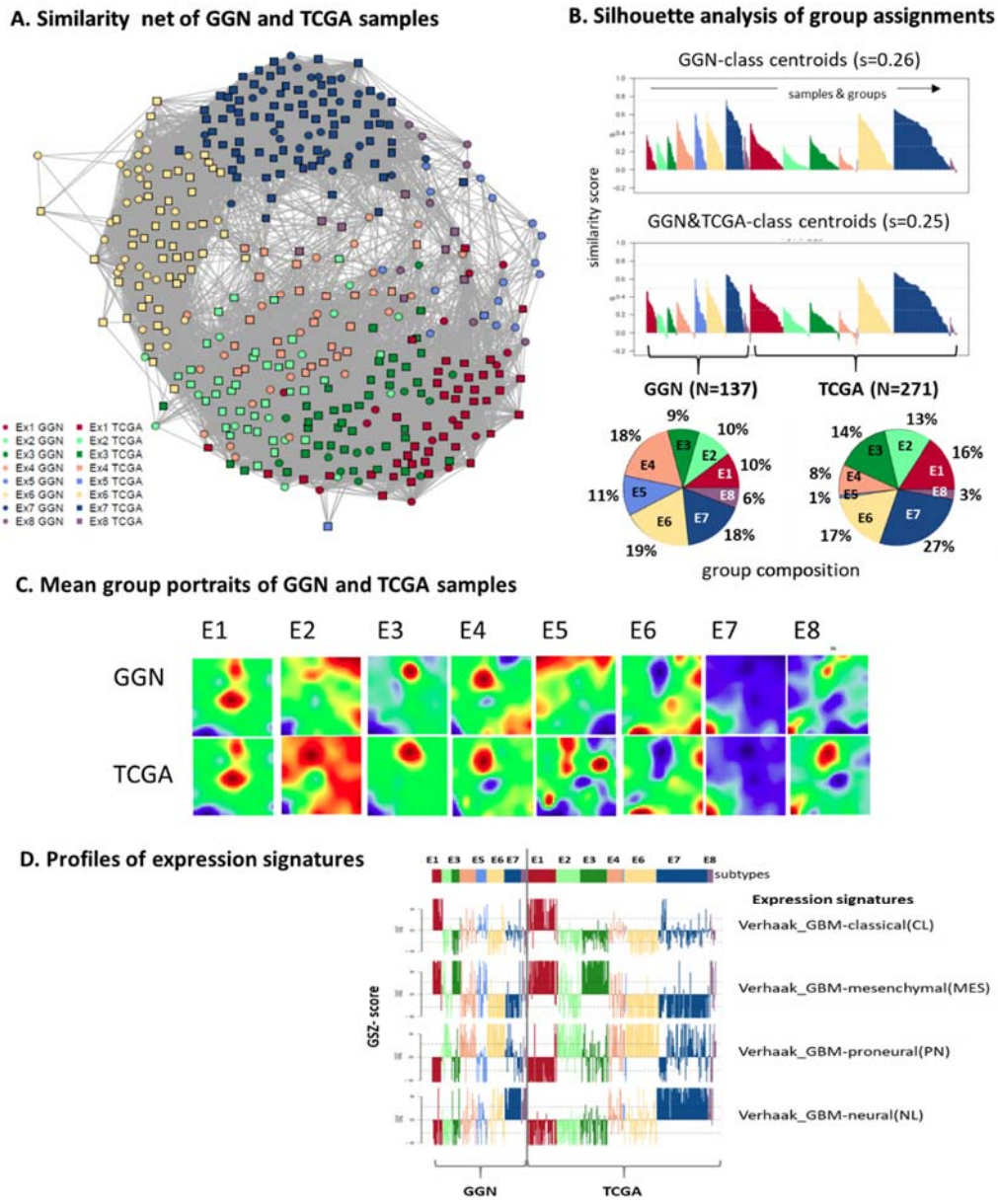


Figure S 8: Validation of the subtyping of the expression landscapes using a TCGA cohort of lower grade (II/III) glioma that comprises matched expression data of 271 tumors and methylation data of 278 tumors (Table S 2). Both, the GGN and TCGA data sets of LGG samples were trained together using a GGN-guided extension SOM-clustering of the TCGA samples [29]. TCGA samples were classified according to closest similarity to the class centroids of the GGN-samples. A) The similarity network confirms that GGN (circles) and TCGA (squares) samples of the same class cluster together and form common data clouds. B) Predominantly positive values of the silhouette cluster scores for GGN and TCGA samples confirm proper classification. Using cluster centroids of the GGN-samples of each class or of both, the GGN and TCGA samples, provided almost identical results. All sample groups except E5 were reproduced in the TCGA data. The compositions of both cohorts regarding the subtypes except E5 resemble each other. C) Mean SOM-group portraits of the GGN- and TCGA-cohorts confirm agreement of the expression landscapes of the subtypes except E5. Details and possible reasons for this fact will be discussed in the main paper. (D) Expression profiles of transcriptional signatures of the Verhaak-GBM classes (see Table S 6) across the E-subtypes assigned to samples of the GGN and TCGA data sets. Both data sets are in strong agreement with subtype related expression. Note also the sharp changes of the expression levels from subtype-to-subtype which confirms proper classification. Interestingly, E3 shows mesenchymal and reduced proneural characteristics in contrast to the other IDH

mutated subtypes what explains the partial similarity between E3 and the IDH-wt E1- groups evident in the heatmap in Figure 1A.

SOM extension method: The SOM extension method [29] aims at adding new, secondary sample data (e.g. TCGA tumors here) to an already existing SOM (e.g. that of the GGN tumors here) and thereby adapting the secondary data to the primary SOM space. This extended SOM is obtained by repeating the training process for the combined data where however only the primary data were used to calculate the intrinsic metrics of the SOM during initiation and adaptation. This 'pickaback' training assigns unchanged metagene values to the primary data and new, adapted metagene data to the secondary data.

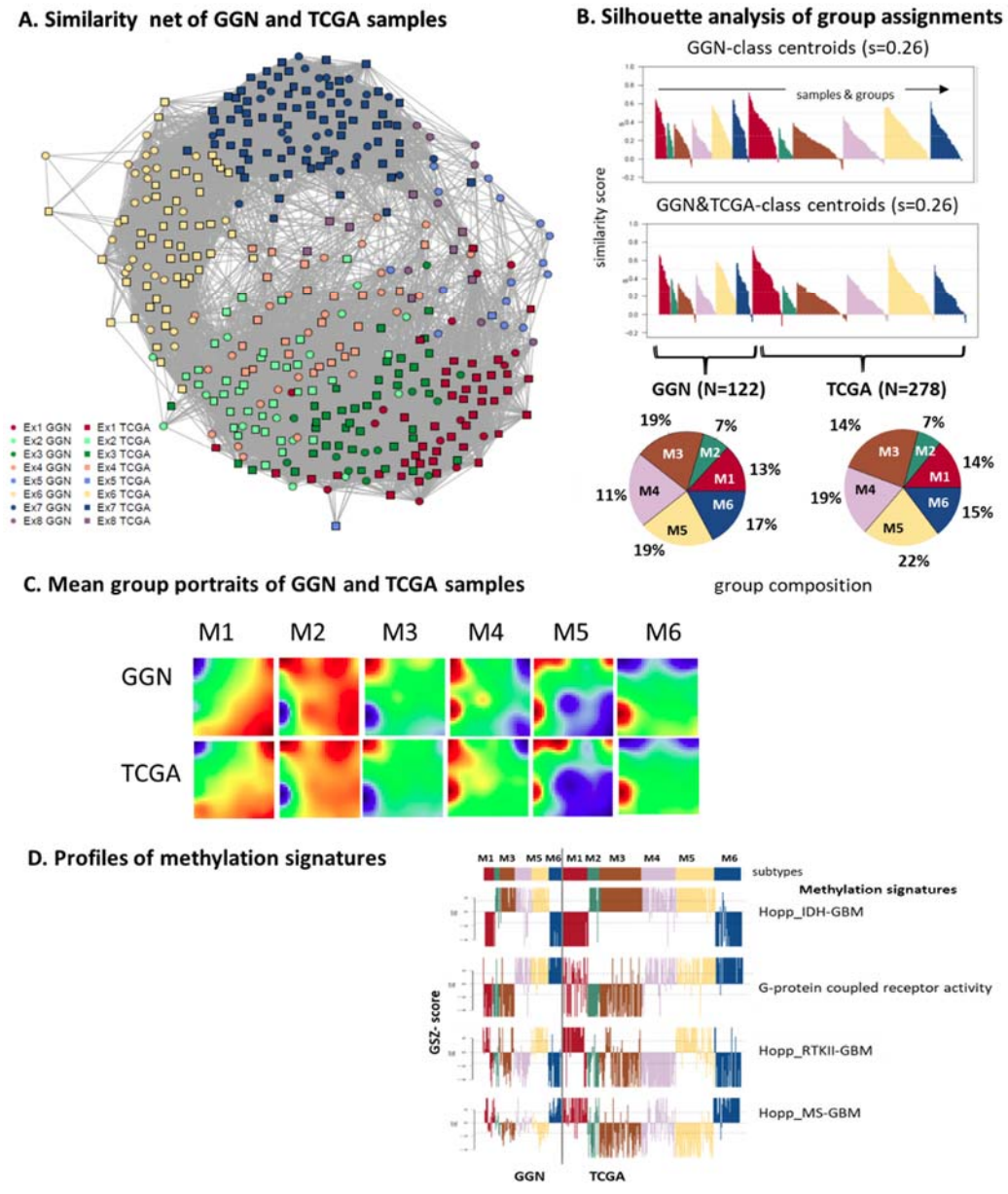


Figure S 9: Validation of the subtyping of the methylation landscapes using the TCGA cohort of lower grade glioma. All methylation subtypes were reproduced in the GGN cohort. See legend of Figure S 9. Also a set of marker CpG-genes published in [20] for classification of LGG well differentiates between the most of the M-groups identified here (Figure S 10). (D) Methylation profiles of methylation signatures

as proposed by Sturm et al. (Table S 6) and of *GPCRs* across the subtypes assigned to samples of the GGN and TCGA data sets. Both data sets are in strong agreement with subtype related methylation patterns, e.g. of the IDH subtype reflecting hypermethylation in IDH-mut tumors of both data sets, but also for the GPCR signature showing hypomethylation in M1 – M3 in GGN and TCGA data as well. Note that the RTKII and MS methylation signatures designed to classify IDH-wt GBM differentiate between IDH-O and IDH-A LGG samples of the GGN and TCGA cohorts as well.

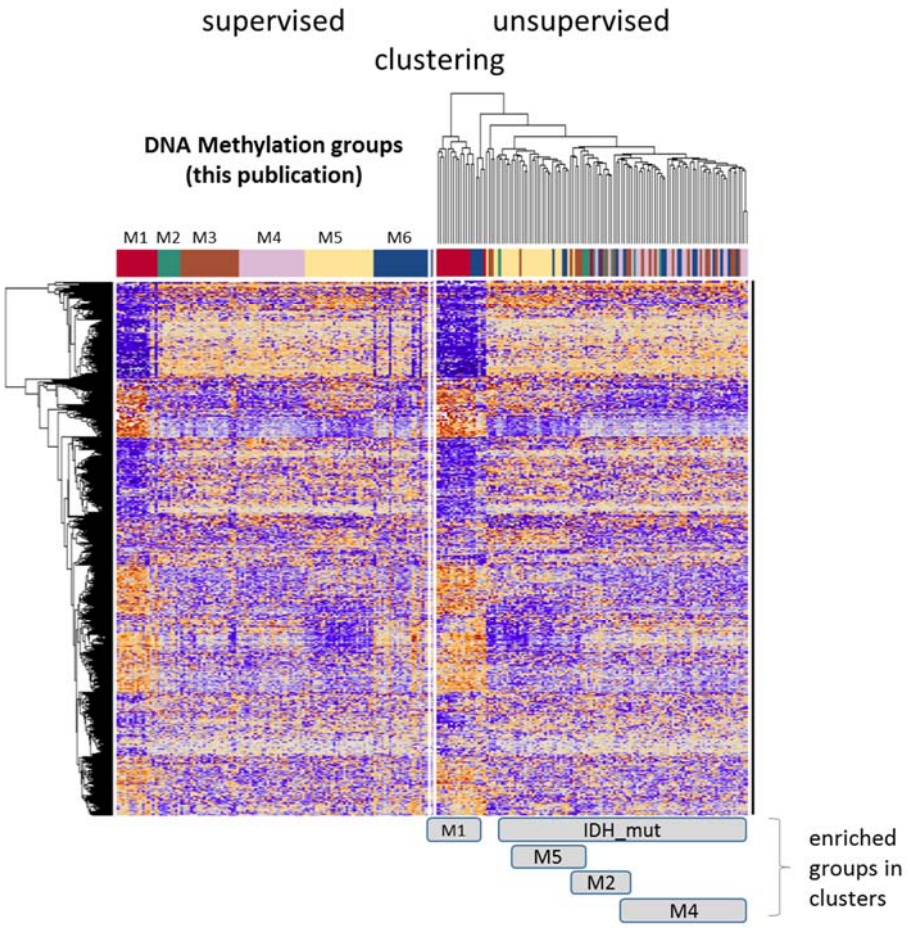


Figure S 10: Supervised and unsupervised clustering of samples using 1308 signature CpG classifying LGG in [20]. Virtually four main unsupervised sample clusters were found corresponding to M1, M5 and to M2 and M4 according to the enrichment of GGN-samples. Samples of the M6 subtype split according to their *IDH* mutation status where *IDH* mutated tumors cluster together with M1.

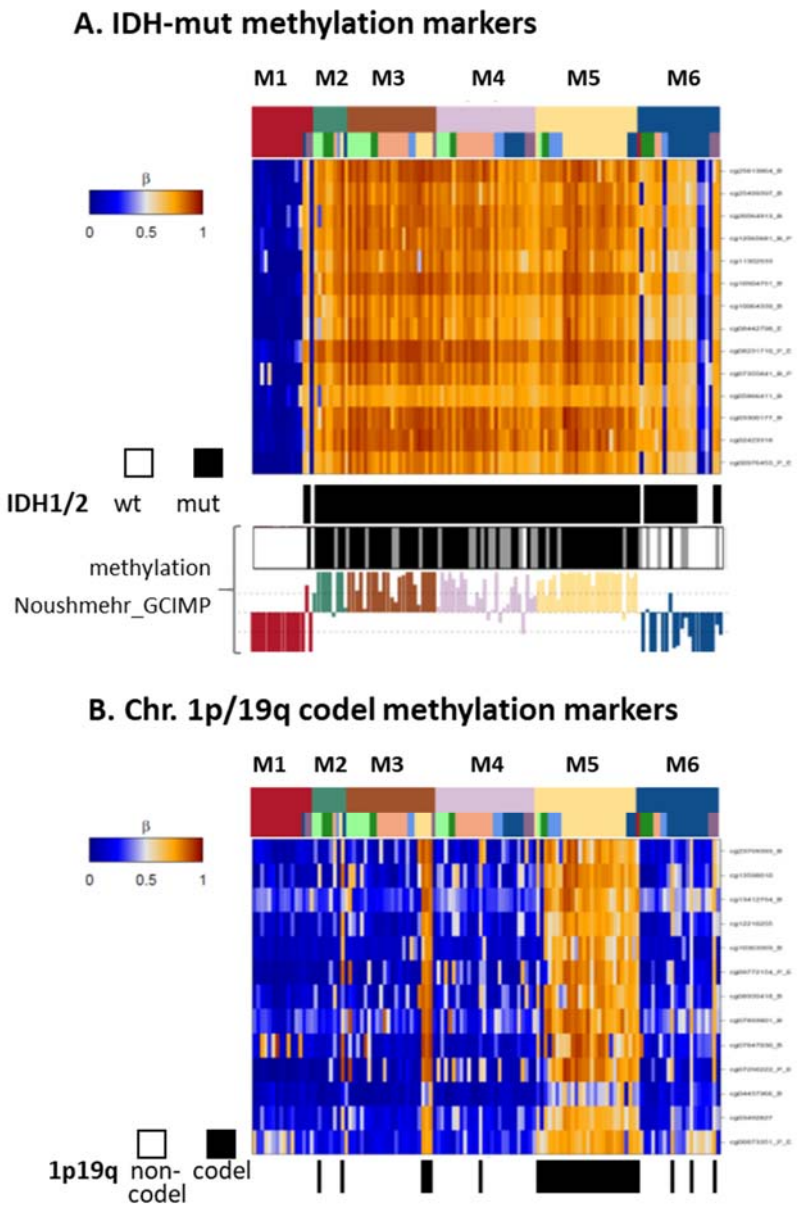


Figure S 11: Heatmaps of marker CpGs for IDH mutated (part A) and Chr 1p/19q codel (B) tumors taken from [37]. Both methylation profiles confirm the genetic data shown as black (mutated/aberrant)/white (non-mutated/intact) barcodes for IDH mutated and Chr 1p/19q tumors, respectively. Also the signature of Noushmehr et al. (Table S 6) which considers hypermethylated and underexpressed genes in GCIMP tumors agrees with the IDH-mut status however with less contrast in the M6 (neural) subtype presumably due to an relatively high content of healthy tissue in M6 samples.

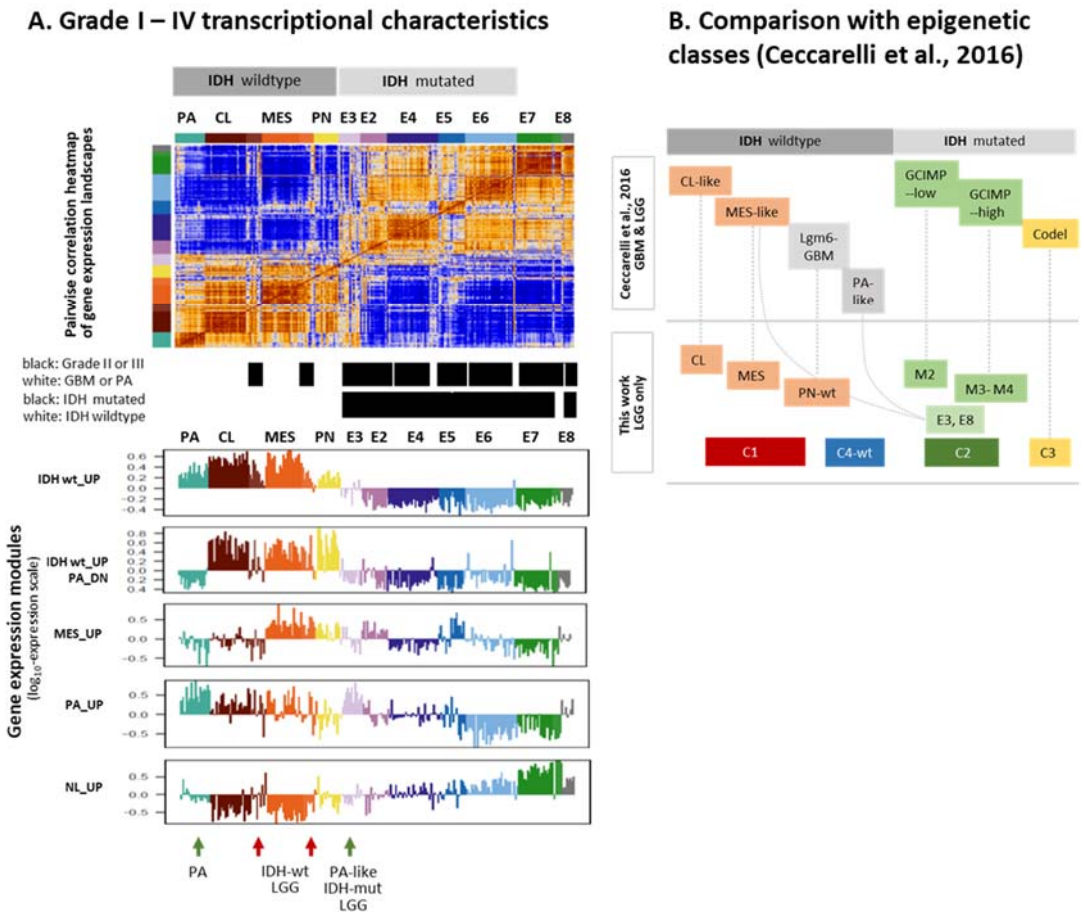


Figure S 12: Comparison of classification of LGG proposed in this work with the epigenetic classes of glioma provided by Ceccarelli and colleagues [20]. (A) Pairwise correlation heatmap of the full GGN-glioma data set, which comprises the LGG studied in this work, 94 grade IV GBM described previously [7] and 16 pilocytic astrocytomas (PA) collected in the GGN. Our LGG data set was extended by GBM and PA to make it comparable with the glioma set of Ceccarelli et al. For a joint subtyping which includes grade I to IV gliomas we classified all IDH-wt LGG to the GBM IDH-wt GBM subtypes used in [7], namely CL (classical), MES (mesenchymal) and PN-wt (proneural) except for the neural ones (E7 and E8). All twelve IDH-mut grade IV GBM were classified into the LGG classes E2 – E8 (see the class label bar and the WHO-grade bar in the figure). Then, the transcriptomic data were clustered into five major modules of co-regulated genes whose profiles were shown below the heatmap. Firstly, we find that expression of E3 resembles that of PA (green arrows) while expression of E2 partly resembles that of MES IDH-wt GBM. Secondly, grade II and III IDH-wt tumors assigned to CL and MES subtypes slightly differ from grade IV IDH-wt by elevated neuronal characteristics (NL_UP module) and reduced IDH-wt characteristics (see red arrows). Note also the continuous decay of the NL_UP signature from E7 to E3/E2 and IDH-wt which roughly associates with increased grading and/or tumorigenicity. (B) The scheme compares our subtypes with the seven classes of Ceccarelli et al. Our CL, MES and PN-wt tumors correspond to the CL-like, MES-like and Lgm6-GBM classes of Ceccarelli et al., respectively. The PA_UP profile in panel A suggests that part of the CL and especially MES tumors resemble PA-characteristics and eventually can be assigned to a PA-like subtype not explicitly considered by us. More interestingly, E3 and, partly, E8 IDH-mut LGG also show PA-resemblance. The GCIMP-low, GCIMP-high and codel classes of Ceccarelli et al. can be assigned to our classes M2 – M5 as indicated by the dashed lines. Note also that M2 shows worse prognosis compared with the other IDH-mut groups in correspondence with the poor prognosis of GCIMP-high tumors while E3 shows relatively good prognosis in correspondence with PA-like tumors of Ceccarelli et al. (see Figure S 6). Note also good agreement between M2 (9% of IDH-mut of grade II-IV in GGN) and GCIMP-low (6% of IDH-mut tumors of grade II-IV in TCGA) regarding the percentage of cases.

Hence, our data support the classification of Ceccarrelli et al. while we, in addition, identified IDH-mut LGG with PA-resemblance.

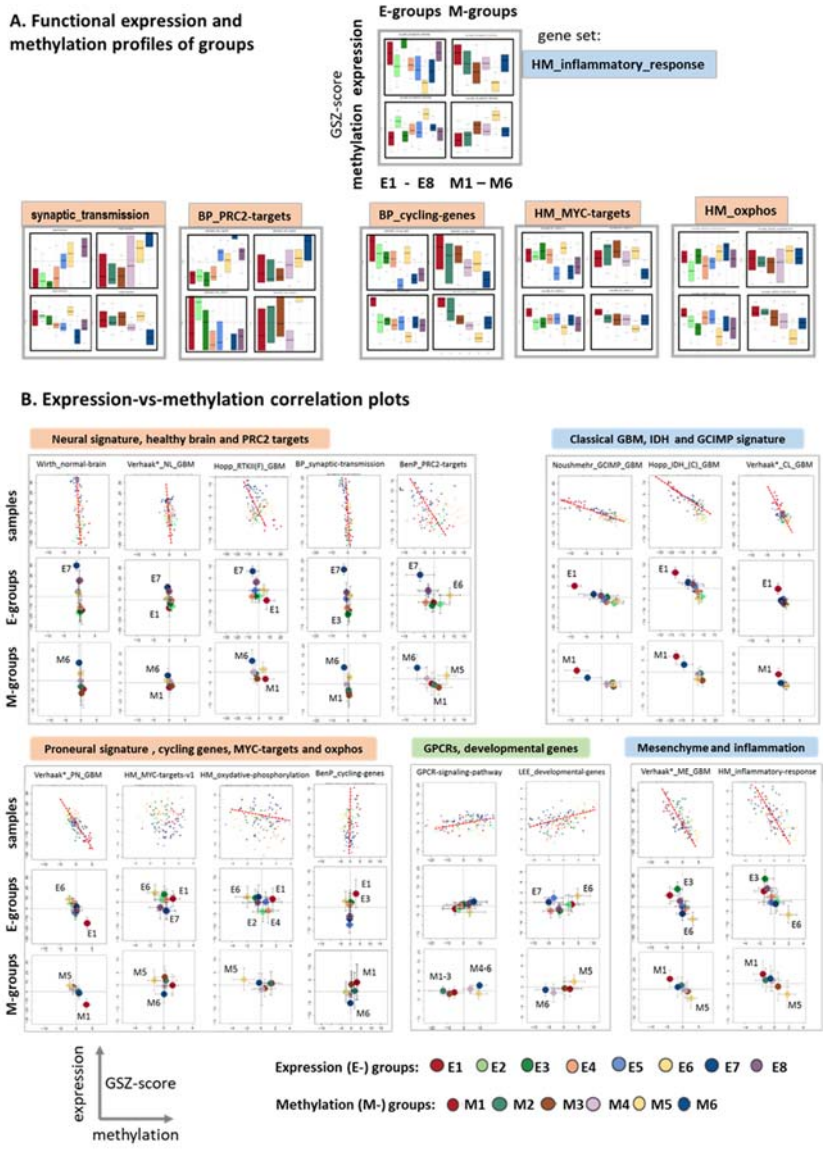


Figure S 13. Functional context of the expression and methylation GSZ-profiles of E- and M-groups: A) Each gene set signature splits into a quadruple of profiles, namely the expression and methylation signatures as revealed by E- and M-groups. Part B) shows the respective expression-versus-methylation correlation plots with sample and group resolution. Overall, the profiles of E- and M-subtypes resemble each other according to their mutual correspondence. The expression and methylation profiles are clearly anti-correlated in most cases which reflects a repressive effect of gene promoter methylation on the expression of the downstream gene on the average. The gene sets refer to different functional categories as indicated in the figure. Interestingly, the expression of G-protein coupled receptors (*GPCR*s) and of developmental genes positively correlate with methylation.

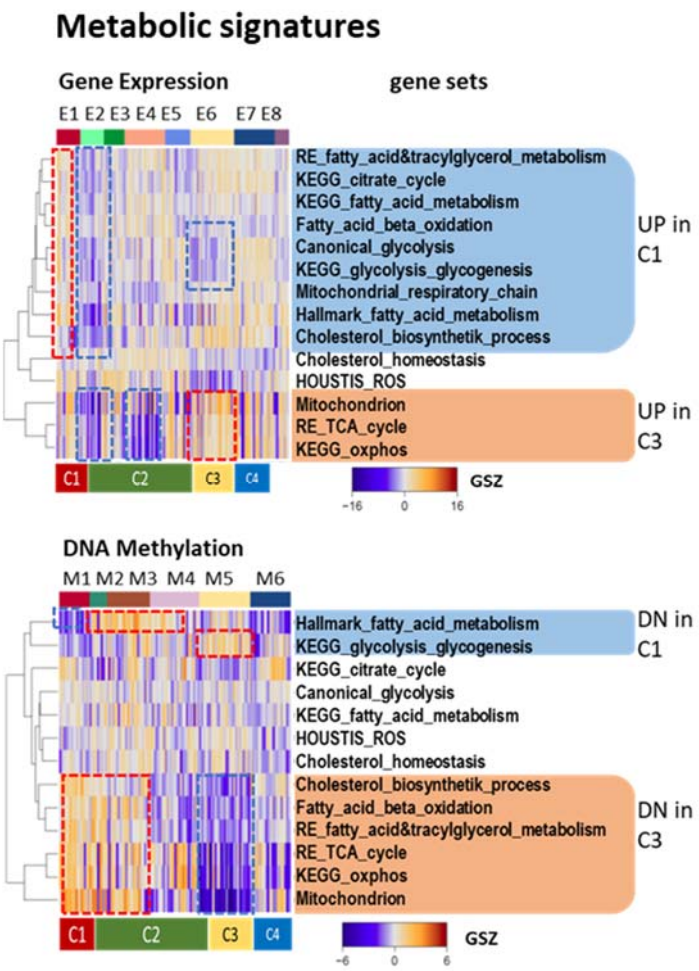
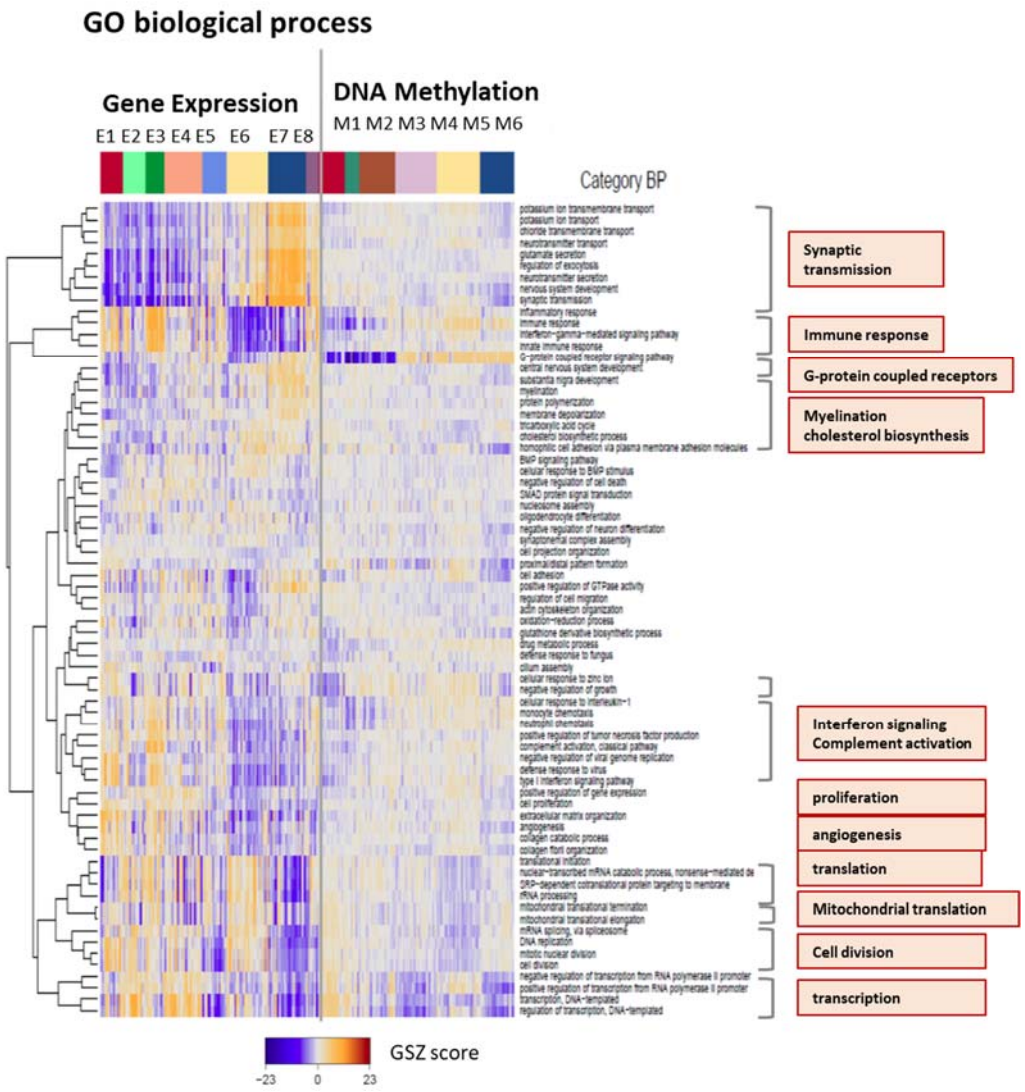


Figure S 14: Heatmaps of selected metabolic gene signatures indicate subtype-specific activation and de-activation patterns.


GO biological process
Gene Expression
E1 E2 E3 E4 E5 E6 E7 E8
DNA Methylation
M1 M2 M3 M4 M5 M6
Category BP
potassium ion transmembrane transport
potassium ion transport
chloride transmembrane transport
neurotransmitter transport
glutamate secretion
regulation of exocytosis
neurotransmitter secretion
nervous system development
synaptic transmission
inflammatory response
immune response
interferon-gamma-mediated signaling pathway
innate immune response
G-protein coupled receptor signaling pathway
central nervous system development
substantia nigra development
myelination
protein polymerization
membrane depolarization
tricarboxylic acid cycle
cholesterol biosynthetic process
homophilic cell adhesion via plasma membrane adhesion molecules
BMP signaling pathway
cellular response to BMP stimulus
negative regulation of cell death
SMAD protein signal transduction
nucleosome assembly
oligodendrocyte differentiation
negative regulation of neuron differentiation
synaptonemal complex assembly
cell projection organization
proximal/distal pattern formation
cell adhesion
positive regulation of GTPase activity
regulation of cell migration
actin cytoskeleton organization
oxidation-reduction process
glutathione derivative biosynthetic process
drug metabolic process
defense response to fungus
cilium assembly
cellular response to zinc ion
negative regulation of growth
cellular response to interleukin-1
monocyte chemotaxis
neutrophil chemotaxis
positive regulation of tumor necrosis factor production
complement activation, classical pathway
negative regulation of viral genome replication
defense response to virus
type I interferon signaling pathway
positive regulation of gene expression
cell proliferation
extracellular matrix organization
angiogenesis
collagen catabolic process
collagen fibril organization
translational initiation
nuclear-transcribed mRNA catabolic process, nonsense-mediated de
SRP-dependent cotranslational protein targeting to membrane
rRNA processing
mitochondrial translational termination
mitochondrial translational elongation
mRNA splicing, via spliceosome
DNA replication
mitotic nuclear division
cell division
negative regulation of transcription from RNA polymerase II promoter
positive regulation of transcription from RNA polymerase II promoter
transcription, DNA-templated
regulation of transcription, DNA-templated
Synaptic transmission
Immune response
G-protein coupled receptors
Myelination
cholesterol biosynthesis
Interferon signaling
Complement activation
proliferation
angiogenesis
translation
Mitochondrial translation
Cell division
transcription
GSZ score
-23
0
23

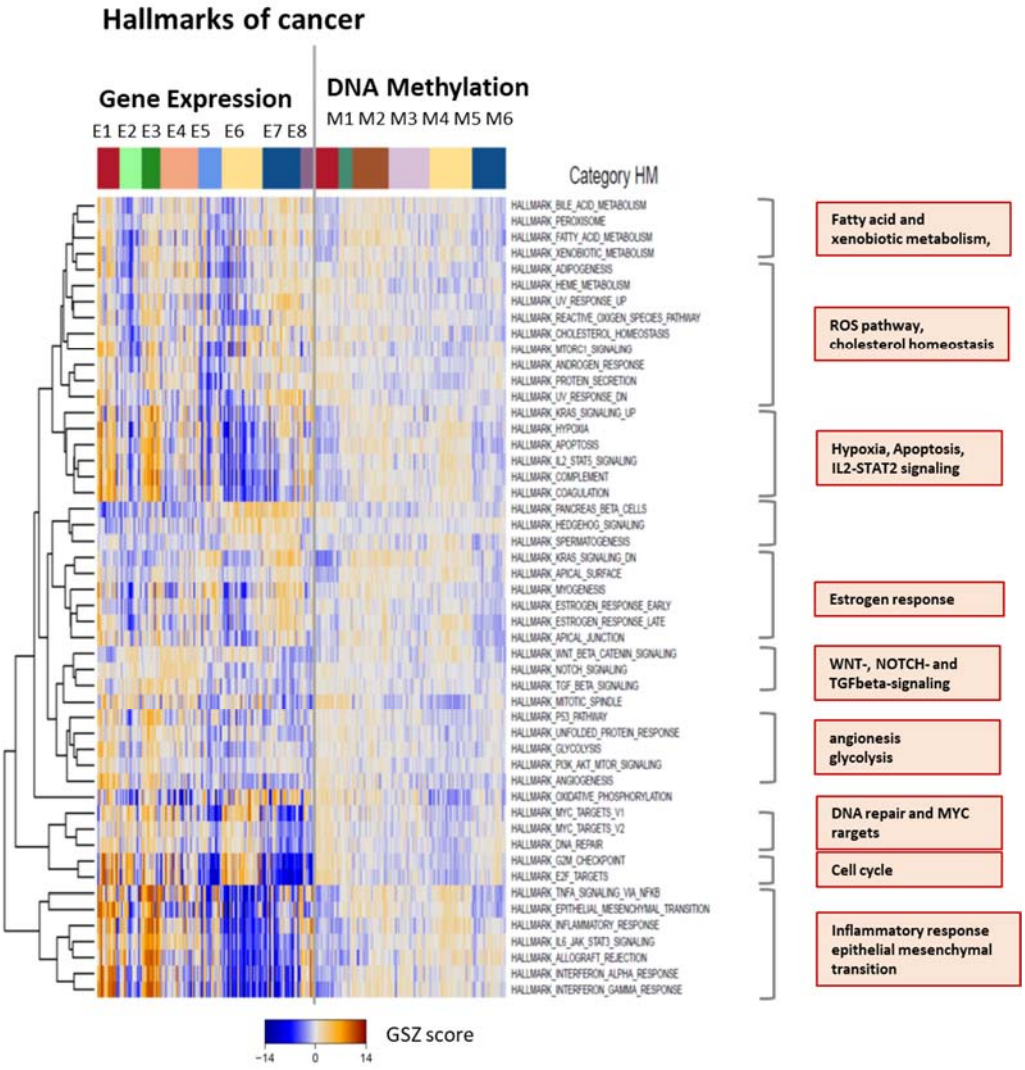


Figure S 15: Gene set analysis of functional signatures taken from the gene ontology category biological process (BP) and from the category hallmarks of cancer (HM) [76].

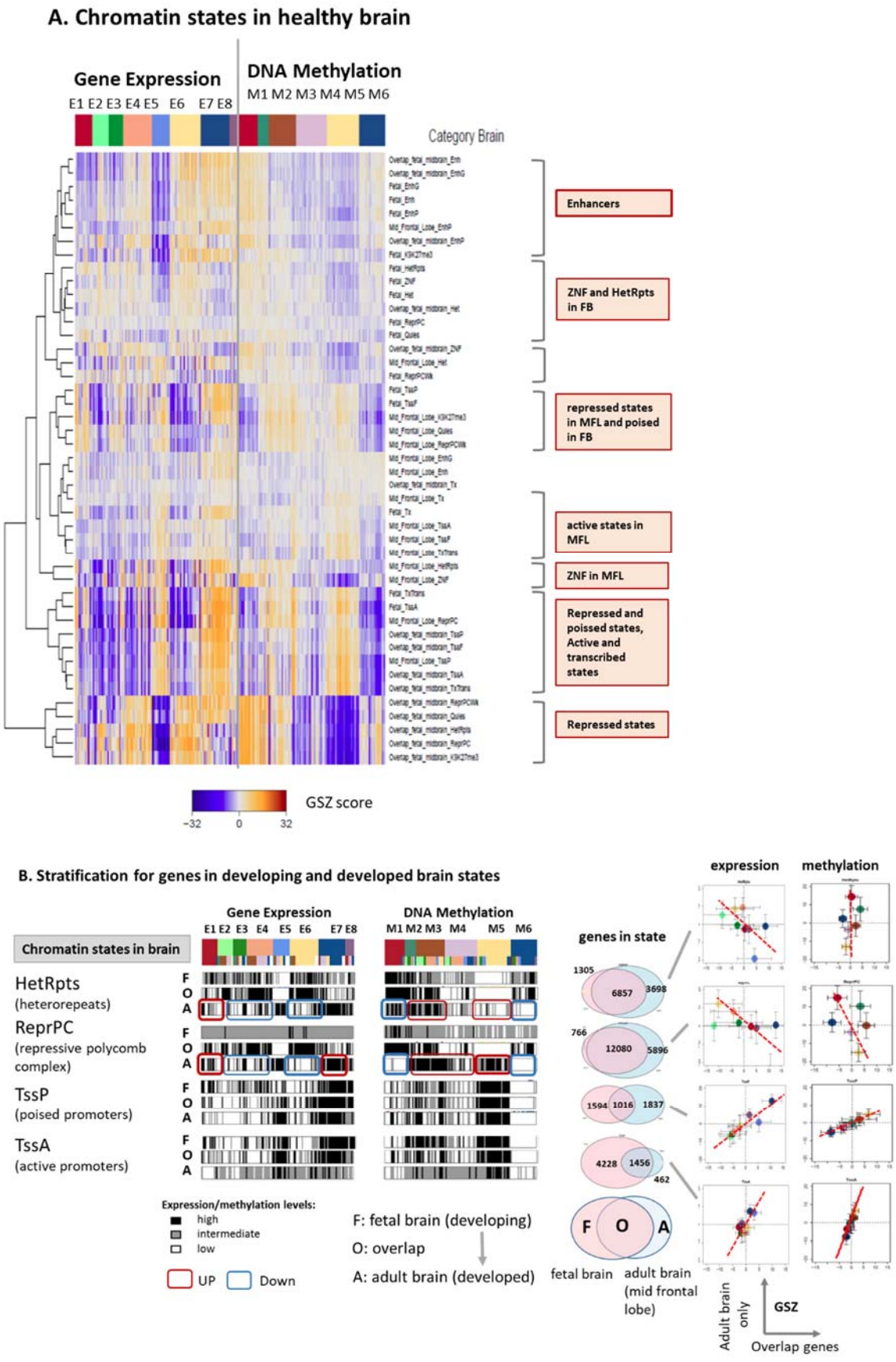

A. Chromatin states in healthy brain
Gene Expression
E1 E2 E3 E4 E5 E6 E7 E8
DNA Methylation
M1 M2 M3 M4 M5 M6
Category Brain
Enhancers
ZNF and HetRpts in FB
repressed states in MFL and poised in FB
active states in MFL
ZNF in MFL
Repressed and poissed states, Active and transcribed states
Repressed states
GSZ score
−32 0 32
B. Stratification for genes in developing and developed brain states
Chromatin states in brain
HetRpts
(heterorepeats)
ReprPC
(repressive polycomb complex)
TssP
(poised promoters)
TssA
(active promoters)
F
O
A
Expression/methylation levels:
high
intermediate
low
UP
Down
F: fetal brain (developing)
O: overlap
A: adult brain (developed)
genes in state
1305
6857
3698
766
12080
5896
1594
1016
1837
4228
1456
462
expression
methylation
fetal brain
adult brain (mid frontal lobe)
Adult brain only
GSZ
Overlap genes

Figure S 16: Gene set analysis of epigenetic promoter states derived from fetal and adult healthy brain [54]. We assigned genes to promoter states as determined in [53] for developed adult brain (midfrontal lobe) and developing fetal brain tissue. Genes were divided into three categories, namely if they were found uniquely either in developing or developed brain (F- and A-genes, respectively); or together in both (overlap O-genes). The latter O-genes seem to have less impact for brain development while the F- and A-genes are assumed to be potentially related to brain development because they change state between developing and developed brain. A.) Overview heatmap of expression and methylation levels of gene promoters in the chromatin states considered (see also the glossary-list given below for assignments). Essentially one finds two major clusters referring predominantly to O-genes either in repressed or in active and poised states (two marked clusters from below). B.) For a more detailed evaluation mean expression and methylation levels of F-, O- and A-genes in selected states are shown as barcode plots. The Venn-diagrams provided their number distributions (only a relatively small number of genes is specifically repressed in the developing brain showing that repression promotes differentiation of brain tissue) and correlation plots between the expression and methylation of O- and A-genes. The patterns observed divide essentially into two types: (i) Genes in TssP and TssA states show similar expression (and methylation) profiles for F-, O- and A-genes giving rise to positive correlation between A- and O-genes. (ii) O- and A-genes in HetRpts and especially ReprPC states change their mean expression and methylation profiles especially in C3 giving rise to more puzzling, virtually negative mutual correlations. In other words, genes in repressed states with impact for brain differentiation become hypermethylated in G-CIMP-subtypes and especially in C3.

The states were defined as follows (http://egg2.wustl.edu/roadmap/web_portal/chr_state_learning.html):

1	TssP	transcription start site (TSS)_poised
2	TssF	TSS_flanking_more_upstream
3	TssA	TSS_active
4	Tx	Transcription
5	EnhG	Transcription Enhancer-like
6	Enh	Enhancer_active_with_weakK4me1_strong_K27ac
7	EnhP	Enhancer_poised
8	ReprPCWk	Repressed_polycomb_weak
9	ReprPC	Repressed_polycomb
10	K9K27me3	H3K9me3_K27me3
11	ZNF	Zinc_finger_genes_H3K36me3_K9me3
12	HetRpts	Heterochromatin_at_repeats
13	Het	Heterochromatin
14	Quies	Quiescent
15	K9acLow	low H3K9ac

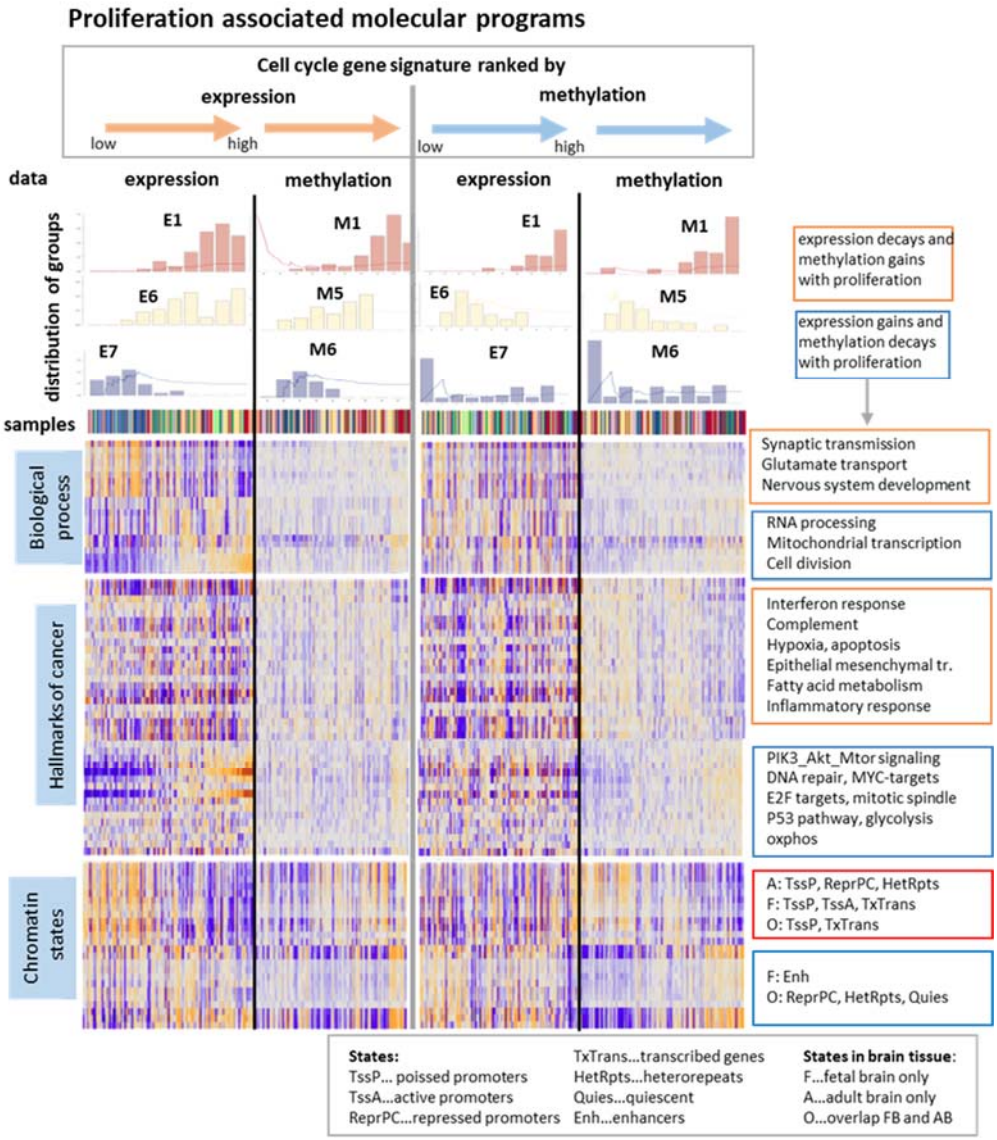


Figure S 17: Proliferation-ranked analysis of gene expression and methylation for selected functional categories and chromatin states: Tumor samples were ranked with increased expression of the gene set 'cell-cycle_literature' using the GSZ-metrics [77] (left part, see red arrows above) and with increased methylation of this set (right part, see blue arrows). The frequency distribution of selected glioma subtypes (part above) and of expression and methylation levels are shown for both rankings. One sees that tumors of the consensus class C1 (E1 and M1) associate with high expression and methylation levels of the cell-cycle related cellular program while C4 (E7 and M6) show low expression and widely distributed methylation levels. Subtype C3-tumors (E6 and M5) associate with low methylation and intermediate to high expression levels for these functions. Hence, methylation shows the opposite trend in C1 and C3 with increasing proliferative activity, namely low methylation of the cell-cycle related genes (cell division, RNA processing and mitochondrial transcription) in C3 and higher levels in C1. Transcriptional activation of cell cycle related biological processes were opposed by the decay of neuronal processes such as synaptic transmission. For methylation one finds similar parallels for processes related to inflammation which associates with the decreased immune response in C3. Interestingly, the two different patterns related to neuronal activity and inflammation on one hand (red frames) and to cell cycle activity (blue frames) associate with different chromatin states in healthy brain, namely first of all poised promoters in fetal and adult brain in the first case, which suggests that these states were affected independently of their impact for brain development. Contrarily, repressed genes in adult, developed brain show antagonistic expression and methylation compared with repressed states observed also in fetal brain. The latter ones show parallels in their methylation and, to a less degree, expression patterns with cell-cycle related processes while the former states associate with neuronal processes.

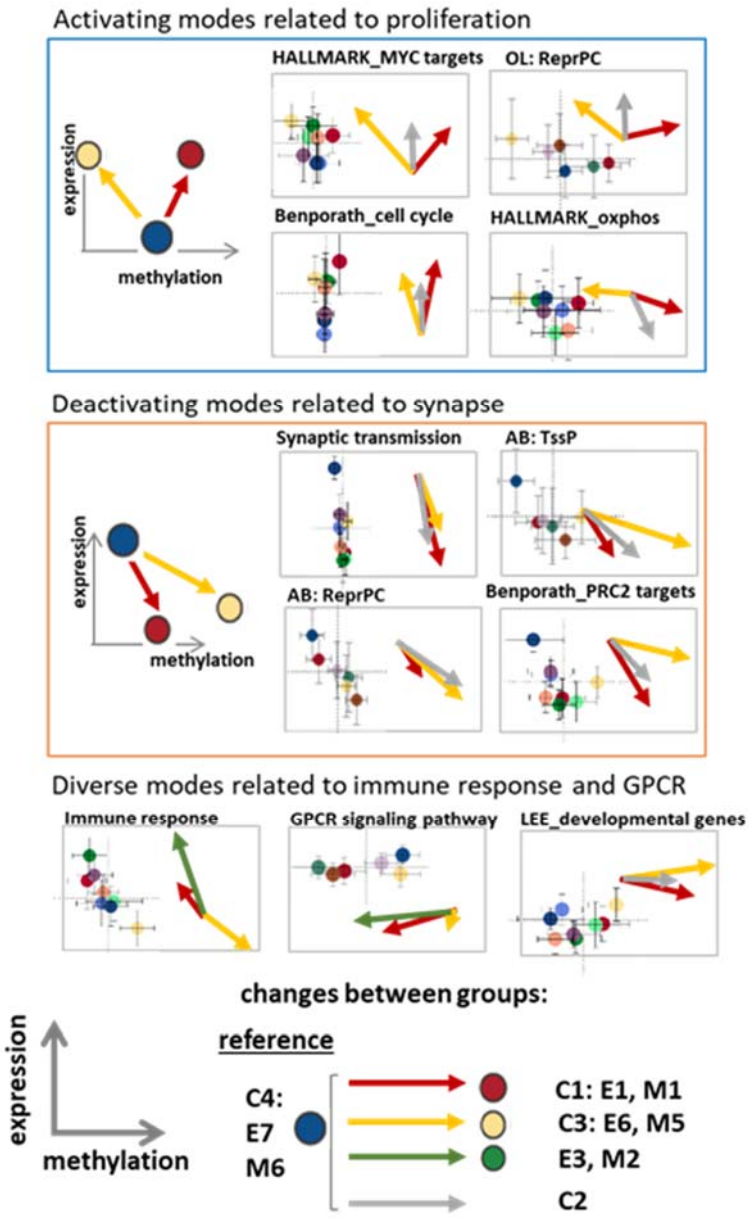
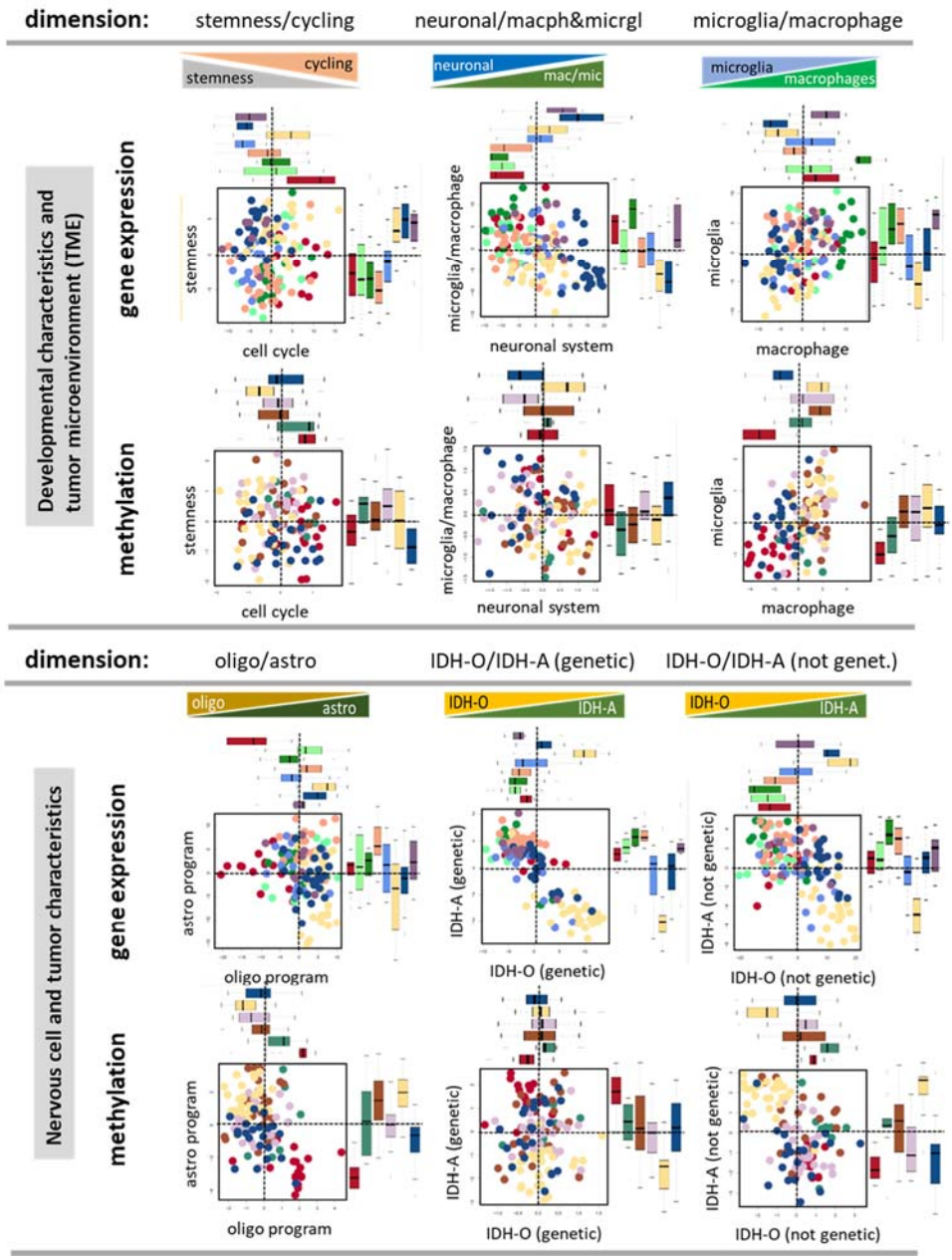


Figure S 18: Combined expression-methylation patterns: For a combined view on expression and methylation changes of selected gene sets we visualize them as arrows pointing from C4 that was chosen as 'brain-like' reference state towards the glioma subtypes C1, C2 and C3 in the expression-versus-methylation biplots. Overall three different combinations were identified: (i) The activating modes were related to proliferation and show increased expression which however associates either with increased (C1) or decreased (C3) methylation reflecting different driving mechanisms. (ii) Deactivating modes combine decreased expression and increased methylation in all subtypes and include functions such as synaptic transmission and a series of epigenetic signatures related to poised and repressed promoters. (iii) Functions related to immune response also show anti-correlated changes between expression and methylation but an activating effect in C1 and especially E3 and deactivating effect in C3. Hence, degeneration of healthy brain functions in all subtypes, activated proliferation in C3 and partly inflammation in E3 seem to be affected by anti-correlated DNA-promoter-methylation changes. Vector-type plots show arrows that point from C4 to C1 (red), C3 (yellow) or M2/E3 (grey) in the expression-vs-methylation plots.

A. Dimensions of glioma heterogeneity: Stratification into subtypes



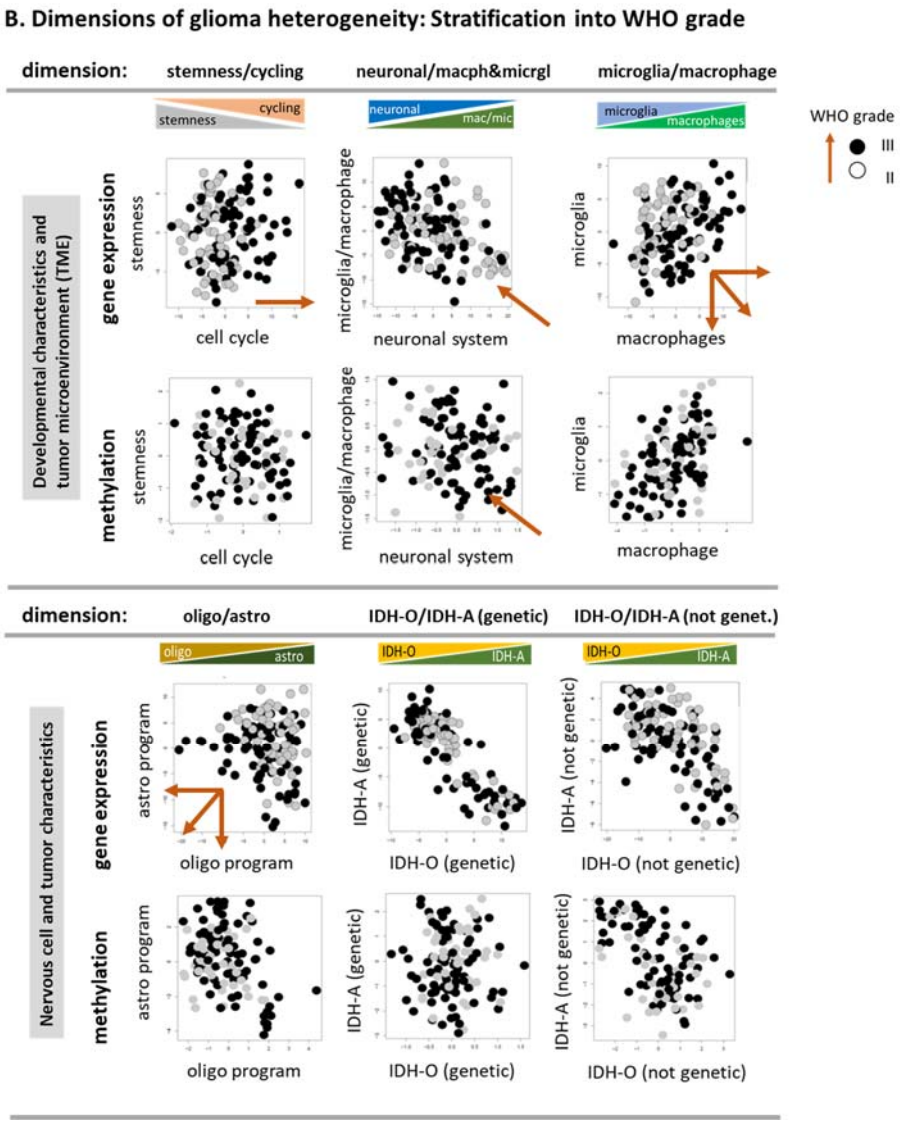


Figure S 19: A) Plots of pairwise combinations of single-cell signatures taken from [56] characterize different dimensions of glioma heterogeneity. The dots are the samples color-coded according to E1 – E8 (expression data) and M1 – M6 (methylation data) that were further summarized into barplots shown at the respective axes of the plots. Firstly, one sees that expression signatures of malignant astrocyte-like (*IDH*-A) and oligodendrocyte-like (*IDH*-O) cells are on highest levels in the C2 and C3 subtypes, respectively, as expected. E3 and E4 show the strongest activity of the *IDH*-A and astro- signatures which suggests a high content of astrocyte-like cells. Secondly, C3 and C4 (neuronal) are almost similar in their expression levels with respect to neuronal, stemness, macrophage signatures while C2 are either reduced or enhanced, respectively. Thirdly, C3 and C4 considerably differ with respect to cell cycle activity where that of C3 is close to that of C2 and C1. Fourth, anti-correlated expression and methylation patterns are found especially for the malignant *IDH*-A and *IDH*-O dimensions suggesting that neoplastic transformations in *IDH*-O and *IDH*-A cells are driven by de-methylation of the respective signature genes. Moreover, *IDH*-O and *IDH*-A expression and methylation levels change antagonistically between C1/C2 and C3 presumably due to anticorrelated amounts of *IDH*-A and *IDH*-O cell. Part B) shows the same plots as in part A however colored according to the WHO-grade (II…grey; III…black). The red arrows visualize trends of increasing WHO grade in the data. Accordingly, cell cycle and microglia/macrophage signatures gain while neuronal, astro- and oligo-signatures decline with increasing grade, on the average.

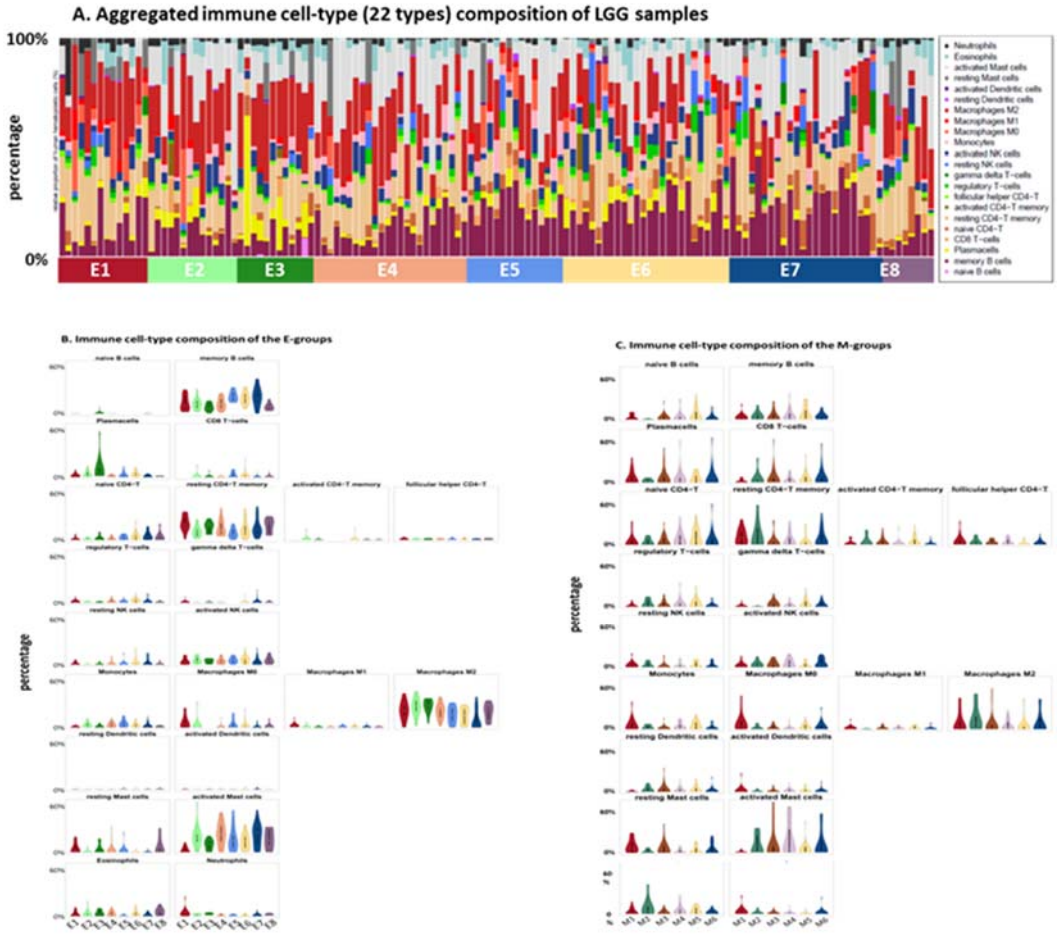


Figure S 20: Digital immune cell-type deconvolution of glioma transcriptomes using CIBERSORT [32]. The program applies support vector regression based on previous knowledge about purified leukocyte expression profiles to estimate the percentage of 22 immune cells (part A) which were further stratified into 11 cell-types (Figure 5B). Part B and C present violin plots of the compositions of the expression and methylation subgroups, respectively. Relative large amounts of macrophages are found in E1 – E3 while B-cells accumulate in E5 – E7 (part A). Interestingly, especially M2-macrophages are enriched in the astrocytic groups E1 – E4 with highest levels in E3 and E2 accumulating higher grade III tumors. M2-macrophages play a pro-tumoral role in brain cancer; they pursue an anti-inflammatory function, promote tissue remodeling and tumor growth [57]. Moreover, high levels of M2-macrophages associate with resistance to radiotherapy in grade IV GBM of the mesenchymal subtype [24]. In contrast, anti-tumoral and pro-inflammatory M1-macrophages are almost absent in all LGG subtypes studied. Large M2-macrophage abundance in astrocytic gliomas is paralleled by relatively large amounts of resting CD4-memory cells, especially in E3 and E1, and by reduced amounts of resting CD8 T-cells. The abundance of tumor infiltrating CD4+ leukocytes in GBM correlates with tumor progression and relates to tumor angiogenesis while it anti-correlates with infiltrating CD8+ leucocytes [62, 63]. We also found that activated mast cells are relatively abundant in virtually all groups of predominantly *IDH*-mutated tumors (E2 – E8), especially in E4, E5 and E7; and on relatively low level in the *IDH*-wt group E1. Mast cells were shown to become recruited and 'educated' by glioma cells in a glioma grade-dependent manner and reduce stemness, decrease proliferation and migration but in turn induce differentiation of glioma cells [64].

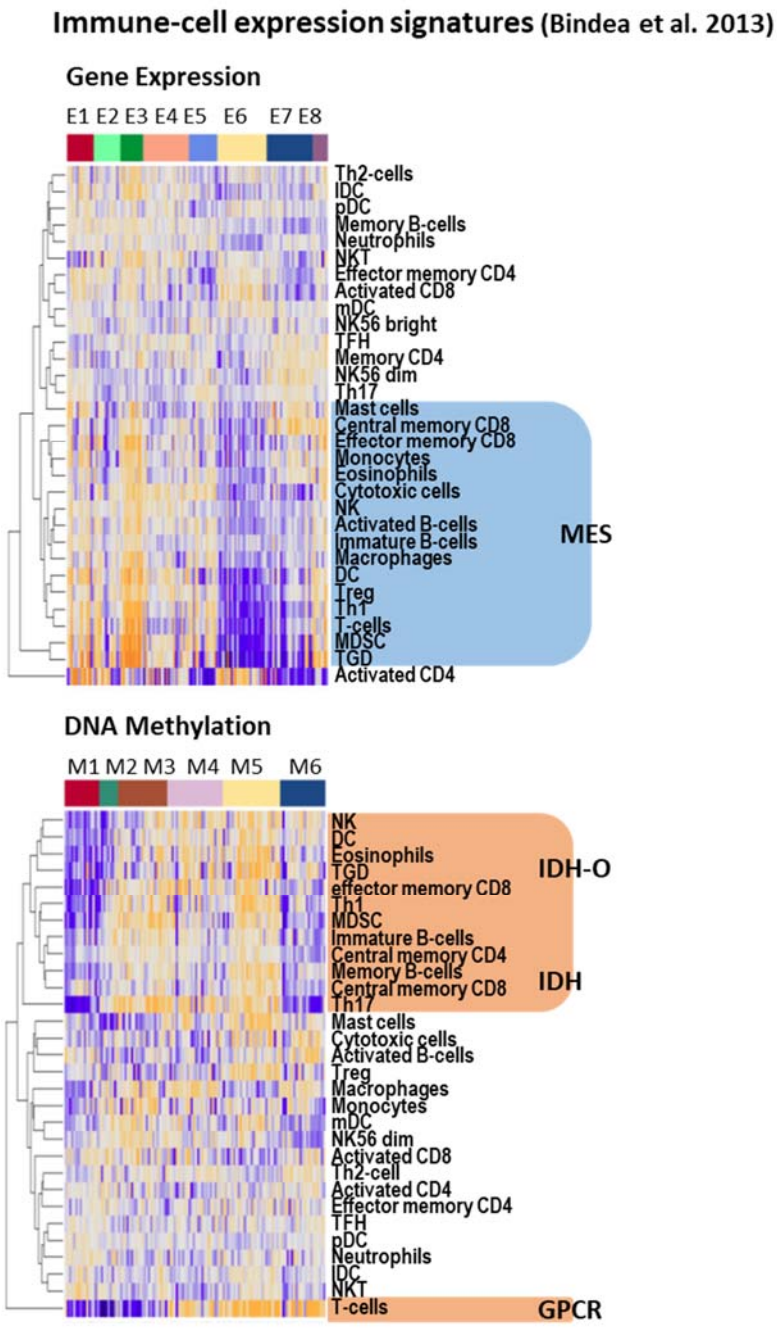


Figure S 21: Immune cell expression signatures taken from [66] confirm the signatures provided by immune cell deconvolution using CIBERSORT (Figure S 20). In contrast they estimate absolute levels of the expression signatures (versus the percentages provided by Cibersort) and they deliver also methylation profiles (part below). A series of immune cell signatures (e.g. natural killer cells (NK), effector and central memory and immature and memory B cells) reveal a G-CIMP/*IDH* and/or *IDH*-O methylation patterns while in gene expression most immune cells show a mesenchymal signature upregulated in higher grade astrocytoma (especially in E3 and E1) that supports the view that mesenchymal gliomas accumulate a series of immune cells. Interestingly, T-cells reveal a *GPCR* methylation signature and a MES expression profile. Cytotoxic cells appear activated in astrocytoma-like gliomas due to demethylation.

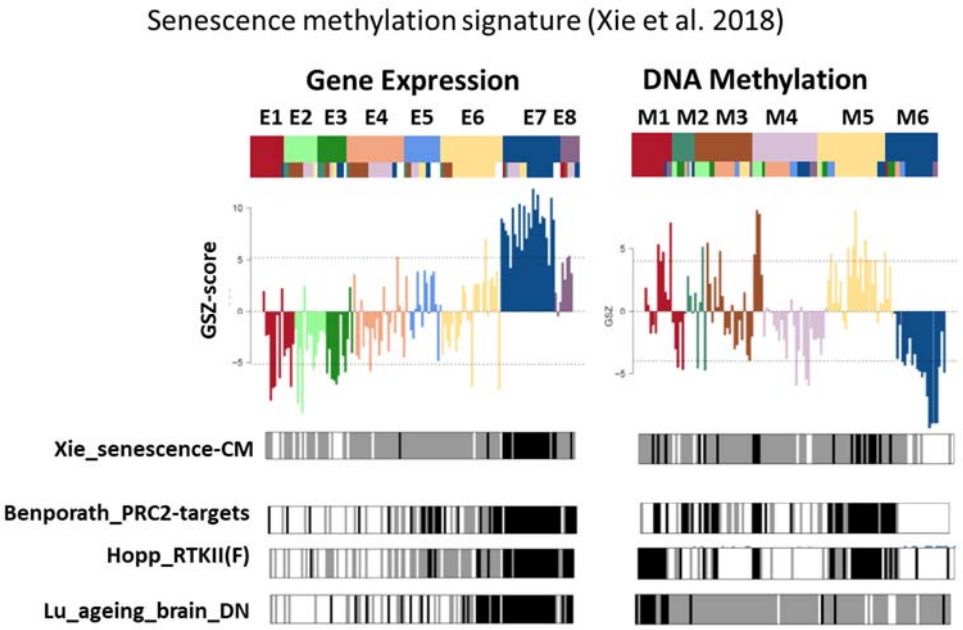


Figure S 22. The senescence methylation signature taken from [68] associates with continuously decaying expression levels of the samples in the E-groups from E7 to E1, i.e. from neuronal via pro-neural towards mesenchymal-like characteristics. This trend suggests increased senescence in parallel with tumor development and formation of a pro-inflammatory microenvironment at later stages in agreement with a recent model of glioma progression [67]. Notably, the expression and methylation profiles of the senescence signature closely resembles the profiles of the PRC2-targets and of the RTKII-methylation (GCIMP-O) signatures (see Figure 2 and Figure 3). Interestingly, the profiles of the expression signatures of the 'ageing brain and of healthy brain functions such as synaptic transmission also resemble that of the senescence signature (Figure 2 and Figure 3) while their methylation profiles differ regarding methylation of C3. It is assumed that Chr. 1p/19q-codeleted gliomas bypass senescence by other mechanisms than Chr. 1p/19q-non-codeleted tumors [67]. Our results indeed indicate that C3-tumors show increased methylation of senescence genes which however result in only moderate transcriptional deactivation compared with C1 and C2 tumors possibly because of activated oxphos-metabolism (Figure S 14) and/or deactivated inflammatory response. The 'commonly methylated' (CM) senescence signature genes become promoter-hypermethylated during aging and tumorigenesis and are thought to mediate retention of proliferating, aging cells and emerge as potential biomarkers of cancer risk [68]. They are enriched for developmental regulators which become suppressed to favor anti-differentiation and self-renewal mechanisms and eventually also for genes regulating biosynthetic and metabolic processes.

Supplementary Tables

Table S 1: List of GGN-patients whose DNA methylation data were included into the study (separate Excel-file).

Table S 2: List of TCGA-samples used for verification analyses (separate Excel-file).

Table S 3: Mutual sample distribution between expression groups defined in this publication and; A) the expression groups presented previously [6] using semi-supervised classification; B) the methylation groups of this work; C) the genomic groups presented in [6] and, D) between the methylation groups of this work and the genomic groups presented in [6]. Increased mutual overlaps are indicated by grey color.

A) Expression groups in this work versus expression groups of Weller et al. (2015) [a]										
		Expression groups (this work)								
		E1	E2	E3	E4	E5	E6	E7	E8	Total
Expression groups (Weller et al., 2015) [b]	gr 1	11								11
	gr 2	3				1	1		1	6
	gr 3							5		5
	gr 4		8	11		1	1		1	22
	gr 5		3		22					25
	gr 6		3			12				15
	gr 7					1	24	1		26
	gr 8			1	2			18	6	27
	Total	14	14	12	24	15	26	24	8	137

[a] Overall one finds almost an one-to-one relation between both classifications for E4 (gr. 5), E5 (gr. 6), E6 (gr. 7). Gr. 4 distributes over E2 and E3 and gr. 8 over E7 and E8 (grey fields). In turn, E1 distribute over gr.1 and 2 while E7 over gr. 3 and gr. 8. Hence, here we present a more granular classification mainly for gr.4 and gr. 8 that were split into two new strata each. On the other hand, *IDH*-wt tumors, especially of gr. 2 were assigned to strata that otherwise contain *IDH*-mut tumors.

[b] g.1- gr.3 contain exclusively *IDH*-wt tumors while gr.4- gr.8 were assigned to *IDH*-mut tumors.

B) Expression groups versus methylation groups										
		Expression groups (this work)								
		E1	E2	E3	E4	E5	E6	E7	E8	Total
Methylation groups (this work)	M 1	13						1	2	16
	M 2		2	3	1	1	1	1		9
	M 3		6	2	8	2	4	5	1	28
	M 4		4	1	10	3			3	21
	M 5		1	2		3	18	3		27
	M 6	1		3	2	2		11	2	21
	Total	14	13	11	21	11	23	21	8	122

C) Expression groups versus genomic groups of Weller et al. (2015) [a]										
		Expression groups (this work)								
		E1	**E2**	**E3**	**E4**	**E5**	**E6**	**E7**	**E8**	**Total**
Genomic gr.(Weller et al., 2015) [b]	**gr I**			2		4	24	6	1	**37**
	gr II		6	2	13	4		6		**31**
	gr III		8	8	11	6	1	7	6	**47**
	gr IV	4				1		4		**9**
	gr V	10					1	1	1	**13**
	Total	**14**	**13**	**11**	**21**	**11**	**23**	**21**	**8**	**122**

[a] gr. IV (*IDH*-wt and Chr7 gains) accumulate in neural E7 while gr. II (*IDH*-mut and Chr7 gains) widely distribute into E2, E4 and E7.

[b] gr.I: *IDH*-mut and Chr.1p/19q codel; gr. II: *IDH*-mut and Chr7gains; gr. III: *IDH*-mut; gr. IV: *IDH*-wt; gr. V: *IDH*-wt and Chr7 gains and Chr10 losses.

D) Methylation groups versus genomic groups of Weller et al. (2015)										
		Expression groups (this work) [a]								
		M 1	**M 2**	**M 3**	**M 4**	**M 5**	**M 6**			**Total**
Genomic groups (Weller et al., 2015)	**gr I**		1	3		25	3			**32**
	gr II		2	10	9		4			**25**
	gr III	2	6	10	17	1	9			**45**
	gr IV	3					4			**7**
	gr V	12					1			**13**
	Total	**17**	**9**	**23**	**26**	**26**	**21**			**122**

[a] gr. II (*IDH*-mut and Chr7 gains) accumulate in M3

Table S 4: Distribution of cases with distinct genetic and clinical characteristics in expression groups ('*' indicates significant enrichment with p<0.05 according to Fisher's exact test)

		Ex1	***Ex2***	***Ex3***	***Ex4***	***Ex5***	***Ex6***	***Ex7***	***Ex8***
Total	number of samples	14	14	12	24	15	26	24	8
		10%	10%	9%	18%	11%	19%	18%	6%
IDH 1/2 status	wildtype	**14***				1	1	5	1
	mutated		14	12	**24***	14	**25***	19	7
Chr1p19q	intact	**14***	12	10	**23***	11	2	18	7
	codel	0	2	2	1	4	**24***	6	1
Chr7	gain	**10***	4	2	8	5	4	4	1
	normal	3	10	10	16	10	22	20	7
Chr10	loss	**10***	2	2	1	4	2	3	2
	normal	4	12	10	**23**	11	24	21	6
Chr1	loss	0	2	2	1	4	24	6	1
	normal	**14***	12	10	**23***	11	2	18	7
Chr19	loss	2	5	4	4	6	**25***	7	2
	normal	**12***	9	8	**20***	9	1	17	6
TERT status	mutated	**8***	1	0	1	4	**24***	5	2
	wildtype	4	**12***	**11***	**23***	10	2	19	6
MGMT status	methylated	6	13	11	20	13	24	18	6
	non methylated	**8***	1	1	4	1	2	6	2
Grade	III	12	10	**12***	10	4	**23***	11	7
	II	2	4	0	**14***	**11***	3	13	1
Histology	O	1	5	7	1	2	**22***	8	4
	A	**13***	9	5	**23***	**13***	4	16	4
Gender	male	5	9	7	15	9	13	18	**8***
	female	**9***	5	5	9	6	13	6	0

Table S 5: Distribution of cases with distinct genetic and clinical characteristics in methylation groups ('*' indicates significant enrichment with p<0.05 acording to Fisher's exact test)

		M1	*M2*	*M3*	*M4*	*M5*	*M6*
Total	number of samples	16	9	23	26	27	21
IDH 1/2 status	wildtype	**14***	0	0	0	0	5
	mutated	2	9	**23***	**26***	**27***	16
Chr1p19q	intact	**16***	7	**20***	**25***	0	18
	codel	0	2	3	1	**27***	3
Chr7	gain	**12***	2	10	5	1	4
	normal	4	7	13	21	**26***	17
Chr10	loss	**11***	2	4	2	0	5
	normal	5	7	19	24	**27***	16
Chr1	loss	0	2	3	1	**27***	3
	normal	**16***	7	**20***	**25***	0	18
Chr19	loss	2	3	9	4	**27***	6
	normal	**14***	6	14	**22***	0	15
TERT status	mutated	**9***	1	4	1	**23***	2
	wildtype	5	8	**19***	**24***	3	**19***
MGMT status	methylated	8	9	20	24	**26***	14
	non methylated	**8***	0	3	2	1	6
Grade	III	13	7	16	14	22	10
	II	3	2	7	12	5	**11***
Histology	O	1	5	7	6	**22***	4
	A	**15***	4	16	20	5	**17***
Gender	male	7	6	11	**22***	16	12
	female	9	3	12	4	11	9

Table S 6: Glioma gene sets

Short name	Comment	Data set	Reference
Expression signatures			
Verhaak* (reanalyzed by Hopp et al.)	Four GBM-subtypes: CL, MES, PN, NL and healthy brain	TCGA, 153 GBM and 10 normal	[78]; [35]
Reifenberger	For GBM-subtypes: CL, MES, PN-wt (*IDH*1/2) and PN-mut	GGN, 94 GBM	[7]
Donson	STS and LTS with inflammatory signature of LTS	26 GBM (mostly pediatric) and AA	[79]
Dong	GBM versus healthy brain	TCGA, 240 GBM and 10 normal	[80]
Weller	Eight expression subtypes E1 – E8	GGN, 137 LGG	[6]
Mukasa	Histological classes and 1p/19q	Grade II-V, 21 gliomas	[81]
Gorovets	Three LGG subtypes: PG (pre-GBM), EPL and NP	101 grade II and III gliomas	[36]
Combined signature			
Noushmehr	G-CIMP signature of genes hypermethylated AND underexpressed in *IDH* mutated GBM	TCGA, 272 GBM	[14]
Methylation signatures			
Laffaire	Diverse methylation signatures	33 LGG and 36 GBM	[11]
Christensen	Gliomas of grade II and III	131 grade II-IV gliomas	[13]
Sturm/Hopp (reanalyzed by Hopp et al.)	Six GBM-subtypes: RTKI, RTKII, MS, *IDH*, G34 and K27	TCGA-GBM including pediatric GBM	[38]; [16]
Shinawi			[82]
Martinez	Methylation signatures	87 GBM	[10]
Wirth	Healthy brain tissue		[27]
Glioma single cell signatures			
Venteicher	Single cell transcriptomics, cell-type signatures of astrocytes, oligodendrocytes, microglia/macrophages and of derived malignant programs		[56]
Glioma single treatment resistance signature			

Segerman	The signature was derived from libraries of glioma-initiating cell clones from glioblastoma (*IDH*-wt) patient samples considering a range of responses to radiation and drugs		[22]

References